Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Linear Algebra and its Applications 395 (2005) 1–82
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

A family of inequalities originating from coding of messages

Joel E. Cohen [a,*], Johannes H.B. Kemperman [b,1], Gheorghe H. Zbăganu [c,2]

[a] *Rockefeller and Columbia Universities, 1230 York Avenue, Box 20, New York, NY 10021-6399, USA*
[b] *Rutgers University, 13 Sandra Road, East Brunswick, NJ 08816, USA*
[c] *Centre for Mathematical Statistics, Casa Academiel, Calea 13 Septembrie Nr. 13, RO-76100 Bucharest 5, Romania*

Received 23 July 2003; accepted 30 May 2004

Submitted by R.A. Brualdi

Abstract

This paper presents 96 new inequalities with common structure, all elementary to state but many not elementary to prove. For example, if n is a positive integer and $a = (a_1, \ldots, a_n)$ and $b = (b_1, \ldots, b_n)$ are arbitrary vectors in $\Re^n_+ = [0, \infty)^n$, and $\rho(m_{i,j})$ is the spectral radius of an $n \times n$ matrix with elements $m_{i,j}$, then

$$\sum_{i,j} \min((a_i a_j), (b_i b_j)) \leqslant \sum_{i,j} \min((a_i b_j), (b_i a_j)),$$

$$\sum_{i,j} \max((a_i + a_j), (b_i + b_j)) \geqslant \sum_{i,j} \max((a_i + b_j), (b_i + a_j)),$$

$$\rho(\min((a_i a_j), (b_i b_j))) \leqslant \rho(\min((a_i b_j), (b_i a_j))),$$

* Corresponding author. Tel.: +1 212 327 8883; fax: +1 212 327 7974.
E-mail addresses: cohen@rockefeller.edu (J.E. Cohen), jkemperman@aol.com (J.H.B. Kemperman), zbagang@tyr.racai.ro, zbagang@csm.ro (G.H. Zbăganu).
[1] Tel.: +1 732 390 4537.
[2] Tel.: +40 1 505755.

0024-3795/$ - see front matter
doi:10.1016/j.laa.2004.05.019

$$\sum_{i,j} \min((a_i a_j), (b_i b_j)) x_i x_j \leqslant \sum_{i,j} \min((a_i b_j), (b_i a_j)) x_i x_j,$$

for all real x_i, $i = 1, \dots, n$,

$$\int\int \log[(f(x) + f(y))(g(x) + g(y))]\, \mathrm{d}\mu(x)\, \mathrm{d}\mu(y)$$
$$\leqslant \int\int \log[(f(x) + g(y))(g(x) + f(y))]\, \mathrm{d}\mu(x)\, \mathrm{d}\mu(y).$$

The second inequality is obtained from the first inequality (which is due to G. Zbăganu [A new inequality with applications in measure and information theories, in: Proceedings of the Romanian Academy, Series A1 (1), 2000, pp. 15–19]) by replacing min with max, and $\times$ with $+$, and by reversing the direction of the inequality. The third inequality is obtained from the first by replacing the summation by the spectral radius. The fourth inequality is obtained from the first by taking each summand as a coefficient in a quadratic form. The fifth inequality is obtained from the first by replacing both outer summations by products, min by $\times$, $\times$ by $+$, and the non-negative vectors a, b by non-negative measurable functions f, g. The proofs of these inequalities are mysteriously diverse.

A nice generalization of the first inequality is proved: Let $*$ be one of the four operations $+$, $\times$, min and max on an appropriate interval J of $\mathfrak{R}$. Let $a, b \in J^n$. Denote by $a * a$ the $n \times n$ matrix $a_{i,j} = a_i * a_j$. Then the matrix $a * a$ is more different from $b * b$ than $a * b$ is from $b * a$. Precisely, if $\|A\| = \sum_{1 \leqslant i,j \leqslant n} |a_{i,j}|$, then $\|a * a - b * b\| \geqslant \|a * b - b * a\|$.

AMS classification: Primary 39B62; Secondary 26D15; 60E15

Keywords: Information theory; Operations research; Brownian bridge; (max, +)-algebra; Quadratic form; Spectral radius

1. Introduction

This paper presents a family of new inequalities, all elementary to state but many not elementary to prove. This introduction explains how these inequalities came to be conjectured, describes some applications in information theory and operations research, and previews the inequalities that will be proved (and disproved).

1.1. Story of this project

In 1999, Zbăganu considered a question in information theory. If one of two messages must be sent over a channel with only two input symbols, A and B, and with n output symbols, $1, \dots, n$, is the chance of error in transmission minimized by sending the first message as AA and the second message as BB, or alternatively by sending the first message as AB and the second message as BA? Zbăganu conjectured that a lower risk that the wrong message will be received is achieved by coding the two messages by the pairs of symbols AA and BB than by the pairs of symbols AB and BA. This result is equivalent to a beautiful inequality: if n is a positive integer and $a = (a_1, \ldots, a_n)$ and $b = (b_1, \ldots, b_n)$ are arbitrary vectors in $\mathfrak{R}^n_+ = [0, \infty)^n$, then

$$\sum_{i,j} \min((a_i a_j), (b_i b_j)) \leqslant \sum_{i,j} \min((a_i b_j), (b_i a_j)). \tag{1.1}$$

Zbăganu proved (1.1) by induction on n. He communicated his results to his colleagues in Bucharest and also (by e-mail in June 1999) to Cohen and Kemperman. The day he received Zbăganu's results, Kemperman found a quick and very different proof of (1.1) involving the covariance function of a Gaussian process closely related to the Brownian bridge. Cohen instead was immediately fascinated by the very simple structure of (1.1). Reading from left to right on each side of the inequality (1.1), one first uses the operator $S =$ addition (summation), next the operator $I =$ minimum, and finally the operator $P =$ multiplication (product). Cohen proposed to call Zbăganu's inequality *SIP*<. He suggested that *SIP*< was one of 64 possible inequalities in which each of S, I, P in Zbăganu's inequality was replaced by each of S, I, P and $A =$ maximum. Kemperman recognized that these four operators could be replaced by commutative operators, leading to a more general question:

Let a and b be arbitrary vectors in $\mathfrak{R}^n$ (possibly required to be non-negative). Let D, E, F be commutative operators (with domain and range to be specified). Assuming compatibility of all operations specified, when is it true that, for all pairs a and b,

$$D(E(F(a_i, a_j), F(b_i, b_j))) \leqslant D(E(F(a_i, b_j), F(a_j, b_i))) \tag{1.2}$$

or else that

$$D(E(F(a_i, a_j), F(b_i, b_j))) \geqslant D(E(F(a_i, b_j), F(a_j, b_i)))? \tag{1.3}$$

Typically, F maps $U \times U$ into V (such as $U = \mathfrak{R}$ and $V = \mathfrak{R}^+$), while E maps $V \times V$ into W (such as $W = \mathfrak{R}$ or $W = \mathfrak{R}^+$) while D operates on $n \times n$ matrices with values in W. The range of D is taken to be some partially ordered set, including possibly all $n \times n$ matrices with the Loewner ordering.

If valid, the inequality (1.2) is denoted by *DEF*<, and (1.3) by *DEF*>, respectively. Except for some equalities, at most one of *DEF*< and *DEF*> will be true. Which of these two has at least a chance to be true can usually be seen from the special case when all elements of a equal one constant and all elements of b equal another. When the inequality is true in general, the direction of the inequality is usually determined by this special case, so one may as well speak briefly of the inequality *DEF*.

Cohen initially considered the 64 inequalities *DEF* when D, E and F are restricted to $\{A = \max, I = \min, S = \text{sum}, P = \text{product}\}$. (At various points, we use different notations for the minimum, so it is useful to be forewarned that $I(x, y) = \min(x, y) = x \wedge y$, for any real x, y. Similarly for the maximum, $A(x, y) = \max(x, y) = x \vee y$.) Each of A, I, S and P can operate on finite sets of any size. Thus $D = S$ (the sum of matrix elements) has a different meaning from $F = S$ (the sum of a pair of numbers), as in the inequality *SAS*. Because A, I, S and P are all associative, *DEF* is true with the equality sign when $E = F$. This observation proved 16 of the 64 inequalities.

Cohen tested numerically the remaining 48 candidate inequalities *DEF* and was very surprised to find that for 46 of them, it was not possible to obtain numerical counterexamples. Kemperman and Zbăganu then undertook the challenge of proving the 46 surviving candidates. This paper reports the proofs of those 46 inequalities, and counterexamples to the other two would-be inequalities.

We also investigated several further extensions of these inequalities. Cohen suggested the case where D is the spectral radius of the non-negative matrix. In this case, we write $D = R$. Zbăganu suggested the case where we replace the summation $D = S$ by a quadratic form. In this case, we write $D = Q$. Each of these two formal mutations of (1.1) led to 16 additional conjectured inequalities, giving a total of $96 = 64 + 16 + 16$ new conjectured inequalities. Zbăganu also suggested the case where we replace the vector pairs a and b by pairs of functions and the summation $D = S$ by an integral.

We believe that these inequalities represent an important new class of inequalities. Despite our efforts, we have not found any universal type of proof. In view of the two exceptional cases, such a universal proof may not exist. On the other hand, if there is a totally new algebraic structure behind many of our results, it might well lead to a better understanding why some results of type *DEF* are true and (a few) others are false.

1.2. Applications

As mentioned above, Zbăganu's inequality (1.1) answered a question in information theory. If a_i represents the probability that the input symbol A is received as the output symbol i and b_j represents the probability that the input symbol B is received as the output symbol j, and if the channel is memoryless so that errors in transmission affect output independently for each input symbol, then the matrix $(a_i a_j)$ is the joint probability distribution of output symbols (i, j) when the input symbols are AA, the matrix $(b_i b_j)$ is the joint probability distribution of output symbols (i, j) when the input symbols are BB, and similarly for the matrices $(a_i b_j)$ and $(b_i a_j)$. The left side of (1.1) measures the similarity between the matrices $(a_i a_j)$ and $(b_i b_j)$ because it takes the value 1 when the matrices are identical and takes the value 0 when the matrices have disjoint support (that is, the elements of one matrix are zero whenever the corresponding elements of the other matrix are positive). Similarly, the right side of (1.1) measures the similarity between the matrices $(a_i b_j)$ and $(b_i a_j)$. Inequality (1.1) shows that a lower risk that the wrong message will be received is achieved by coding the two messages by the pairs of symbols AA and BB than by the pairs of symbols AB and BA. (For teachers, the lesson here may be that if you are trying to teach your students one of two messages, it is better to convey the message twice in the same way than to convey it once in each of two different ways; but this application should not be taken too seriously.)

Generalizations of (1.1) were suggested by generalizations of matrix multiplication important in operations research, including manufacturing theory and routing theory [4,3,1] (and references cited in these sources). If $U = (u_{i,j})_{i,j=1,\ldots,n}$ and $V = (v_{i,j})_{i,j=1,\ldots,n}$ are any two real $n \times n$ matrices, then conventionally $(U \times V)_{i,j} = \sum_k u_{i,k} \times v_{k,j}$. The (max, plus) algebra defines a generalized matrix product $\otimes$ in which the binary operation $\times$ on scalars is replaced by $+$ and the summation of n scalars is replaced by max: $(U \otimes V)_{i,j} = \max_k(u_{i,k} + v_{k,j})$. The (max, times), (min, plus) and (min, times) generalizations of conventional matrix multiplication are defined similarly. These definitions suggested replacing each of the three operations in (1.1) (addition, min, and multiplication) by each of the four operations, min, max, addition and multiplication.

For example,

$$\sum_{i,j} \max((a_i + a_j), (b_i + b_j)) \geqslant \sum_{i,j} \max((a_i + b_j), (b_i + a_j)) \tag{1.4}$$

is obtained from (1.1) by replacing min with max, and $\times$ with $+$, and by reversing the direction of the inequality. This formula has a natural interpretation in the design of a manufacturing process. Suppose a product has two necessary components, components 1 and 2. Suppose these components are manufactured in parallel. Each component requires a process of two steps, steps 1 and 2. Two machines called A and B can be arranged in one of two manufacturing configurations. In configuration I, component 1 passes through machine A in step 1 and again through machine A in step 2 while component 2 passes through machine B in both steps 1 and 2. In the alternative configuration II, component 1 passes through machine A in step 1 and through machine B in step 2 while component 2 passes through machine B in step 1 and through machine A in step 2. The product is completed when both components have completed both steps. Which manufacturing configuration, I or II, has a shorter average time to produce a product? The time that each machine requires to complete a step depends on the environment in the factory (for example, the temperature or the voltage). Let us suppose that at each step the environment may be in one of n possible states, $i = 1, \ldots, n$, and that these states are equally likely and independent between steps 1 and 2, although identical for both machines at each step. If the environment is in state i at step 1, machine A requires time a_i, and machine B requires time b_i to complete step 1; and exactly the same is true at step 2. Thus if the environment is in state i at step 1 and in state j at step 2 (which will occur with probability $1/n^2$), and if component 1 passes through machine A at step

1 and through machine B at step 2 (as in configuration II), then the time required to make component 1 is $a_i + b_j$, the time required to make component 2 is $b_i + a_j$, and the time required to complete the product is $\max((a_i + b_j), (b_i + a_j))$. If both sides of (1.4) are multiplied by $1/n^2$, then the left side represents the average production time in configuration I while the right side represents the average production time in configuration II. The inequality (1.4) shows that configuration II is preferable to configuration I because it has shorter average production time. The assumption in this example that each state of the environment is equally likely can be replaced by arbitrary probabilities for each environmental state, using the extension to quadratic forms that is described below.

In another example, suppose a factory located at X has 2 suppliers of a hazardous raw material. These suppliers are located at V and Z. The raw material is trucked from V to W in one day, transferred to a fresh truck and trucked from W to X in a second day; and likewise from Z to Y in one day, and then in a fresh truck from Y to X in a second day. The factory uses two trucking companies, A and B, and for legal reasons is obliged to use both companies every day. (The raw material is highly sensitive and the government does not permit the factory to be dependent on a single trucker.) The factory can use plan I or plan II to ship the material. In plan I, company A operates from V to W and from W to X, while company B operates from Z to Y and from Y to X. In plan II, company A operates from V to W and from Y to X, while company B operates from Z to Y and from W to X. The capacity of the trucks operated by both companies depends on the road conditions, which are affected by weather, landslides and forest fires. On any given day, both trucking companies experience the same road conditions. Suppose that under condition $i = 1, \dots, n$, the maximum capacity of the trucks available from company A (or B) is a_i tons (or b_i tons, respectively). If conditions are in state i on the first day and in state j on the second day, then, under plan I, company A can ship $\min(a_i, a_j)$ tons of the material from V to X and company B can ship $\min(b_i, b_j)$ tons from Z to X, so the factory in X can receive $\min(a_i, a_j) + \min(b_i, b_j)$ tons. Under plan II, if conditions are in state i on the first day and in state j on the second day, then the factory can get $\min(a_i, b_j)$ tons of the material from V via W and $\min(b_i, a_j)$ tons from Z via Y, so the factory in X can receive $\min(a_i, b_j) + \min(b_i, a_j)$ tons. Under the worst combination of circumstances (i, j), the factory can count on receiving $\min_{i,j}(\min(a_i, a_j) + \min(b_i, b_j))$ tons under plan I and $\min_{i,j}(\min(a_i, b_j) + \min(b_i, a_j))$ tons under plan II. Inequality *ISI*$>$ in Table 1 tells the factory that plan I assures at least as great a supply of the raw material as plan II. Inequality *ASI*$>$ shows that the maximum possible delivery under plan I is at least as great as that under plan II. If the n conditions are equally likely and independent from one day to the next, then inequality *SSI*$>$ guarantees the company that plan I has at least as great an average delivery of the material as plan II. If condition i occurs with probability p_i and independently from day to day, then *QSI*$>$ guarantees that $\sum_{i,j}(\min(a_i, a_j) + \min(b_i, b_j))p_i p_j \geqslant \sum_{i,j}(\min(a_i, b_j) + \min(b_i, a_j))p_i p_j$, i.e., plan I has a better average delivery rate than plan II.

Table 1 Inequalities of the form *DEF*< or *DEF*>, where $D, E, F \in \{A, I, P, S\}$, excluding the 16 cases *DEF*= when $E = F$. See footnote

DEF	Explicit form and generalizations (when possible)	Proof
IIP<	$\bigwedge_{i,j}((a_i a_j) \wedge (b_i b_j)) \leqslant \bigwedge_{i,j}((a_i b_j) \wedge (b_i a_j))$ $\bigwedge_{x,y}((f(x)f(y)) \wedge (g(x)g(y))) \leqslant \bigwedge_{x,y}((f(x)g(y)) \wedge (f(y)g(x)))$	Easy, Section 3
IIS<	$\bigwedge_{i,j}((a_i + a_j) \wedge (b_i + b_j)) \leqslant \bigwedge_{i,j}((a_i + b_j) \wedge (b_i + a_j))$ $\bigwedge_{x,y}((f(x) + f(y)) \wedge (g(x) + g(y))) \leqslant \bigwedge_{x,y}((f(x) + g(y)) \wedge (f(y) + g(x)))$	Easy, Section 3
IIA<	$\bigwedge_{i,j}((a_i \vee a_j) \wedge (b_i \vee b_j)) \leqslant \bigwedge_{i,j}((a_i \vee b_j) \wedge (b_i \vee a_j))$ $\bigwedge_{x,y}((f(x) \vee f(y)) \wedge (g(x) \vee g(y))) \leqslant \bigwedge_{x,y}((f(x) \vee g(y)) \wedge (f(y) \vee g(x)))$	Easy, Section 3
IPI>	$\bigwedge_{i,j}((a_i \wedge a_j)(b_i \wedge b_j)) \geqslant \bigwedge_{i,j}((a_i \wedge b_j)(b_i \wedge a_j))$ $\bigwedge_{x,y}((f(x) \wedge f(y))(g(x) \wedge g(y))) \geqslant \bigwedge_{x,y}((f(x) \wedge g(y))(f(y) \wedge g(x)))$	Easy, Section 3
IPS<	$\bigwedge_{i,j}((a_i + a_j)(b_i + b_j)) \leqslant \bigwedge_{i,j}((a_i + b_j)(b_i + a_j))$ $\bigwedge_{x,y}((f(x) + f(y))(g(x) + g(y))) \leqslant \bigwedge_{x,y}((f(x) + g(y))(f(y) + g(x)))$	Easy, Section 3
IPA<	$\bigwedge_{i,j}((a_i \vee a_j)(b_i \vee b_j)) \leqslant \bigwedge_{i,j}((a_i \vee b_j)(b_i \vee a_j))$ $\bigwedge_{x,y}((f(x) \vee f(y))(g(x) \vee g(y))) \leqslant \bigwedge_{x,y}((f(x) \vee g(y))(f(y) \vee g(x)))$	Easy, Section 3
ISI>	$\bigwedge_{i,j}((a_i \wedge a_j) + (b_i \wedge b_j)) \geqslant \bigwedge_{i,j}((a_i \wedge b_j) + (b_i \wedge a_j))$ $\bigwedge_{x,y}((f(x) \wedge f(y)) + (g(x) \wedge g(y))) \geqslant \bigwedge_{x,y}((f(x) \wedge g(y)) + (f(y) \wedge g(x)))$	Easy, Section 3
ISP>	$\bigwedge_{i,j}((a_i a_j) + (b_i b_j)) \geqslant \bigwedge_{i,j}((a_i b_j) + (b_i a_j))$ $\bigwedge_{x,y}((f(x)f(y)) + (g(x)g(y))) \geqslant \bigwedge_{x,y}((f(x)g(y)) + (f(y)g(x)))$	Easy, Section 3
ISA<	$\bigwedge_{i,j}((a_i \vee a_j) + (b_i \vee b_j)) \leqslant \bigwedge_{i,j}((a_i \vee b_j) + (b_i \vee a_j))$ $\bigwedge_{x,y}((f(x) \vee f(y)) + (g(x) \vee g(y))) \leqslant \bigwedge_{x,y}((f(x) \vee g(y)) + (f(y) \vee g(x)))$	Easy, Section 3
IAI>	$\bigwedge_{i,j}((a_i \wedge a_j) \vee (b_i \wedge b_j)) \geqslant \bigwedge_{i,j}((a_i \wedge b_j) \vee (b_i \wedge a_j))$ $\bigwedge_{x,y}((f(x) \wedge f(y)) \vee (g(x) \wedge g(y))) \geqslant \bigwedge_{x,y}((f(x) \wedge g(y)) \vee (f(y) \wedge g(x)))$	Easy, Section 3
IAP>	$\bigwedge_{i,j}((a_i a_j) \vee (b_i b_j)) \geqslant \bigwedge_{i,j}((a_i b_j) \vee (b_i a_j))$ $\bigwedge_{x,y}((f(x)f(y)) \vee (g(x)g(y))) \geqslant \bigwedge_{x,y}((f(x)g(y)) \vee (f(y)g(x)))$	Easy, Section 3
IAS>	$\bigwedge_{i,j}((a_i + a_j) \vee (b_i + b_j)) \geqslant \bigwedge_{i,j}((a_i + b_j) \vee (b_i + a_j))$ $\bigwedge_{x,y}((f(x) + f(y)) \vee (g(x) + g(y))) \geqslant \bigwedge_{x,y}((f(x) + g(y)) \vee (f(y) + g(x)))$	Easy, Section 3
	For the *PEF* inequalities, μ is a positive measure	
PIP<	$\prod_{i,j}((a_i a_j) \wedge (b_i b_j)) \leqslant \prod_{i,j}((a_i b_j) \wedge (b_i a_j))$ $\int\int \log[(f(x)f(y)) \wedge (g(x)g(y))]\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) \leqslant \int\int \log[(f(x)g(y)) \wedge (g(x)f(y))]\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)$ $E(\log(f(X)f(Y) \wedge g(X)g(Y))) \leqslant E(\log(f(X)g(Y) \wedge g(X)f(Y)))$	Is *SIS*

(*continued on next page*)

Table 1 (*continued*)

DEF	Explicit form and generalizations (when possible)	Proof
PIS<	$\prod_{i,j}((a_i + a_j) \wedge (b_i + b_j)) \leqslant \prod_{i,j}((a_i + b_j) \wedge (b_i + a_j))$ $\int\int \log[(f(x) + f(y)) \wedge (g(x) + g(y))]\, \mathrm{d}\mu(x)\, \mathrm{d}\mu(y) \leqslant$ $\int\int \log[(f(x) + g(y)) \wedge (g(x) + f(y))]\, \mathrm{d}\mu(x)\, \mathrm{d}\mu(y)$ $E(\log((f(X) + f(Y)) \wedge (g(X) + g(Y)))) \leqslant E(\log((f(X) +$ $g(Y)) \wedge (g(X) + f(Y))))$	Theorem 6.11
PIA<	$\prod_{i,j}((a_i \vee a_j) \wedge (b_i \vee b_j)) \leqslant \prod_{i,j}((a_i \vee b_j) \wedge (b_i \vee a_j))$ $\int\int \log[(f(x) \vee f(y)) \wedge (g(x) \vee g(y))]\, \mathrm{d}\mu(x)\, \mathrm{d}\mu(y) \leqslant$ $\int\int \log[(f(x) \vee g(y)) \wedge (g(x) \vee f(y))]\, \mathrm{d}\mu(x)\, \mathrm{d}\mu(y)$ $E(\log((f(X) \vee f(Y)) \wedge (g(X) \vee g(Y)))) \leqslant E(\log((f(X) \vee$ $g(Y)) \wedge (g(X) \vee f(Y))))$	Is *SIA*
PPI>	$\prod_{i,j}((a_i \wedge a_j)(b_i \wedge b_j)) \geqslant \prod_{i,j}((a_i \wedge b_j)(b_i \wedge a_j))$ $\int\int \log[(f(x) \wedge f(y))(g(x) \wedge g(y))]\, \mathrm{d}\mu(x)\, \mathrm{d}\mu(y) \geqslant$ $\int\int \log[(f(x) \wedge g(y))(g(x) \wedge f(y))]\, \mathrm{d}\mu(x)\, \mathrm{d}\mu(y)$ $E(\log((f(X) \wedge f(Y))(g(X) \wedge g(Y)))) \geqslant E(\log((f(X) \wedge$ $g(Y))(g(X) \wedge f(Y))))$	Is *SSI*
PPS<	$\prod_{i,j}((a_i + a_j)(b_i + b_j)) \leqslant \prod_{i,j}((a_i + b_j)(b_i + a_j))$ $\int\int \log[(f(x) + f(y))(g(x) + g(y))]\, \mathrm{d}\mu(x)\, \mathrm{d}\mu(y) \leqslant$ $\int\int \log[(f(x) + g(y))(g(x) + f(y))]\, \mathrm{d}\mu(x)\, \mathrm{d}\mu(y)$ $E(\log((f(X) + f(Y))(g(X) + g(Y)))) \leqslant E(\log((f(X) +$ $g(Y))(g(X) + f(Y))))$	Corollary 4.10
PPA<	$\prod_{i,j}((a_i \vee a_j)(b_i \vee b_j)) \leqslant \prod_{i,j}((a_i \vee b_j)(b_i \vee a_j))$ $\int\int \log[(f(x) \vee f(y))(g(x) \vee g(y))]\, \mathrm{d}\mu(x)\, \mathrm{d}\mu(y) \leqslant$ $\int\int \log[(f(x) \vee g(y))(g(x) \vee f(y))]\, \mathrm{d}\mu(x)\, \mathrm{d}\mu(y)$ $E(\log((f(X) \vee f(Y))(g(X) \vee g(Y)))) \leqslant E(\log((f(X) \vee$ $g(Y))(g(X) \vee f(Y))))$	Is *SSA* or Corollary 4.10
PSI>	$\prod_{i,j}((a_i \wedge a_j) + (b_i \wedge b_j)) \geqslant \prod_{i,j}((a_i \wedge b_j) + (b_i \wedge a_j))$	False!, True for $n = 2$
PSP>	$\prod_{i,j}((a_i a_j) + (b_i b_j)) \geqslant \prod_{i,j}((a_i b_j) + (b_i a_j))$ $\int\int \log[(f(x)f(y)) + (g(x)g(y))]\, \mathrm{d}\mu(x)\, \mathrm{d}\mu(y) \geqslant$ $\int\int \log[(f(x)g(y)) + (g(x)f(y))]\, \mathrm{d}\mu(x)\, \mathrm{d}\mu(y)$ $E(\log((f(X)f(Y)) + (g(X)g(Y)))) \geqslant E(\log((f(X)g(Y)) +$ $(g(X)f(Y))))$	Theorem 6.1
PSA<	$\prod_{i,j}((a_i \vee a_j) + (b_i \vee b_j)) \leqslant \prod_{i,j}((a_i \vee b_j) + (b_i \vee a_j))$ $\int\int \log[(f(x) \vee f(y)) + (g(x) \vee g(y))]\, \mathrm{d}\mu(x)\, \mathrm{d}\mu(y) \leqslant$ $\int\int \log[(f(x) \vee g(y)) + (g(x) \vee f(y))]\, \mathrm{d}\mu(x)\, \mathrm{d}\mu(y)$ $E(\log((f(X) \vee f(Y)) + (g(X) \vee g(Y)))) \leqslant E(\log(f(X) \vee$ $g(Y) + g(X) \vee f(Y)))$	Implied by *GSA* (Corollary 5.9)
PAI>	$\prod_{i,j}((a_i \wedge a_j) \vee (b_i \wedge b_j)) \geqslant \prod_{i,j}((a_i \wedge b_j) \vee (b_i \wedge a_j))$ $\int\int \log[(f(x) \wedge f(y)) \vee (g(x) \wedge g(y))]\, \mathrm{d}\mu(x)\, \mathrm{d}\mu(y) \geqslant$ $\int\int \log[(f(x) \wedge g(y)) \vee (g(x) \wedge f(y))]\, \mathrm{d}\mu(x)\, \mathrm{d}\mu(y)$ $E(\log((f(X) \wedge f(Y)) \vee (g(X) \wedge g(Y)))) \geqslant E(\log((f(X) \wedge$ $g(Y)) \vee (g(X) \wedge f(Y))))$	Is *SAI*
PAP>	$\prod_{i,j}((a_i a_j) \vee (b_i b_j)) \geqslant \prod_{i,j}((a_i b_j) \vee (b_i a_j))$ $\int\int \log[(f(x)f(y)) \vee (g(x)g(y))]\, \mathrm{d}\mu(x)\, \mathrm{d}\mu(y) \geqslant$ $\int\int \log[(f(x)g(y)) \vee (g(x)f(y))]\, \mathrm{d}\mu(x)\, \mathrm{d}\mu(y)$ $E(\log((f(X)f(Y)) \vee (g(X)g(Y)))) \geqslant E(\log((f(X)g(Y)) \vee$ $(g(X)f(Y))))$	Is *SAS*

Table 1 (*continued*)

DEF	Explicit form and generalizations (when possible)	Proof
PAS>	$\prod_{i,j}((a_i+a_j)\vee(b_i+b_j)) \geqslant \prod_{i,j}((a_i+b_j)\vee(b_i+a_j))$ $\int\int \log[(f(x)+f(y))\vee(g(x)+g(y))]\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) \geqslant$ $\int\int \log[(f(x)+g(y))\vee(g(x)+f(y))]\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)$ $E(\log((f(X)+f(Y))\vee(g(X)+g(Y)))) \geqslant E(\log((f(X)+g(Y))\vee(g(X)+f(Y))))$	Corollary 6.8
	For the *SEF* inequalities, μ is a bounded signed measure and $x\in\Re^n$	
SIP<	$\sum_{i,j}((a_ia_j)\wedge(b_ib_j))x_ix_j \leqslant \sum_{i,j}((a_ib_j)\wedge(b_ia_j))x_ix_j$ $\int\int[(f(x)f(y))\wedge(g(x)g(y))]\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) \leqslant$ $\int\int[(f(x)g(y))\wedge(g(x)f(y))]\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)$ $E(f(X)f(Y)\wedge g(X)g(Y)) \leqslant E(f(X)g(Y)\wedge g(X)f(Y))$	Theorem 5.3, Brownian bridge
SIS<	$\sum_{i,j}((a_i+a_j)\wedge(b_i+b_j))x_ix_j \leqslant \sum_{i,j}((a_i+b_j)\wedge(b_i+a_j))x_ix_j$ $\int\int[(f(x)+f(y))\wedge(g(x)+g(y))]\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) \leqslant$ $\int\int[(f(x)+g(y))\wedge(g(x)+f(y))]\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)$ $E((f(X)+f(Y))\wedge(g(X)+g(Y))) \leqslant E((f(X)+g(Y))\wedge(g(X)+f(Y)))$	Theorem 5.5
SIA<	$\sum_{i,j}((a_i\vee a_j)\wedge(b_i\vee b_j))x_ix_j \leqslant \sum_{i,j}((a_i\vee b_j)\wedge(b_i\vee a_j))x_ix_j$ $\int\int[(f(x)\vee f(y))\wedge(g(x)\vee g(y))]\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) \leqslant$ $\int\int[(f(x)\vee g(y))\wedge(g(x)\vee f(y))]\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)$ $E((f(X)\vee f(Y))\wedge(g(X)\vee g(Y))) \leqslant E((f(X)\vee g(Y))\wedge(g(X)\vee f(Y)))$	Theorem 5.6
SPI>	$\sum_{i,j}((a_i\wedge a_j)(b_i\wedge b_j))x_ix_j \geqslant \sum_{i,j}((a_i\wedge b_j)(b_i\wedge a_j))x_ix_j$ $\int\int[(f(x)\wedge f(y))(g(x)\wedge g(y))]\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) \geqslant$ $\int\int[(f(x)\wedge g(y))(g(x)\wedge f(y))]\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)$ $E((f(X)\wedge f(Y))(g(X)\wedge g(Y))) \geqslant E((f(X)\wedge g(Y))(g(X)\wedge f(Y)))$	Theorem 5.12 Induction on $\|X\|$
SPS<	$\sum_{i,j}((a_i+a_j)(b_i+b_j))x_ix_j \leqslant \sum_{i,j}((a_i+b_j)(b_i+a_j))x_ix_j$ $\int\int[(f(x)+f(y))(g(x)+g(y))]\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) \leqslant$ $\int\int[(f(x)+g(y))(g(x)+f(y))]\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)$ $E((f(X)+f(Y))(g(X)+g(Y))) \leqslant E((f(X)+g(Y))(g(X)+f(Y)))$	Easy
SPA<	$\sum_{i,j}((a_i\vee a_j)(b_i\vee b_j)) \leqslant \sum_{i,j}((a_i\vee b_j)(b_i\vee a_j))$	False!, True for $n=2$
SSI>	$\sum_{i,j}((a_i\wedge a_j)+(b_i\wedge b_j))x_ix_j \geqslant \sum_{i,j}((a_i\wedge b_j)+(b_i\wedge a_j))x_ix_j$ $\int\int[(f(x)\wedge f(y))+(g(x)\wedge g(y))]\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) \geqslant$ $\int\int[(f(x)\wedge g(y))+(g(x)\wedge f(y))]\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)$ $E((f(X)\wedge f(Y))+(g(X)\wedge g(Y))) \geqslant E((f(X)\wedge g(Y))+(g(X)\wedge f(Y)))$	Theorem 5.11
SSP>	$\sum_{i,j}((a_ia_j)+(b_ib_j))x_ix_j \geqslant \sum_{i,j}((a_ib_j)+(b_ia_j))x_ix_j$ $\int\int[(f(x)f(y))+(g(x)g(y))]\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) \geqslant$ $\int\int[(f(x)g(y))+(g(x)f(y))]\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)$ $E((f(X)f(Y))+(g(X)g(Y))) \geqslant E((f(X)g(Y))+(g(X)f(Y)))$	Easy

(*continued on next page*)

Table 1 (*continued*)

DEF	Explicit form and generalizations (when possible)	Proof
SSA<	$\sum_{i,j}((a_i \vee a_j) + (b_i \vee b_j))x_i x_j \leqslant \sum_{i,j}((a_i \vee b_j) + (b_i \vee a_j))x_i x_j$ $\int\int[(f(x) \vee f(y)) + (g(x) \vee g(y))]\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) \leqslant \int\int[(f(x) \vee g(y)) + (g(x) \vee f(y))]\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)$ $E((f(X) \vee f(Y)) + (g(X) \vee g(Y))) \leqslant E((f(X) \vee g(Y)) + (g(X) \vee f(Y)))$	Theorem 5.7
SAI>	$\sum_{i,j}((a_i \wedge a_j) \vee (b_i \wedge b_j))x_i x_j \geqslant \sum_{i,j}((a_i \wedge b_j) \vee (b_i \wedge a_j))x_i x_j$ $\int\int[(f(x) \wedge f(y)) \vee (g(x) \wedge g(y))]\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) \geqslant \int\int[(f(x) \wedge g(y)) \vee (g(x) \wedge f(y))]\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)$ $E((f(X) \wedge f(Y)) \vee (g(X) \wedge g(Y))) \geqslant E((f(X) \wedge g(Y)) \vee (g(X) \wedge f(Y)))$	Implied by *SIA*
SAP>	$\sum_{i,j}((a_i a_j) \vee (b_i b_j))x_i x_j \geqslant \sum_{i,j}((a_i b_j) \vee (b_i a_j))x_i x_j$ $\int\int[(f(x)f(y)) \vee (g(x)g(y))]\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) \geqslant \int\int[(f(x)g(y)) \vee (g(x)f(y))]\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)$ $E((f(X)f(Y)) \vee (g(X)g(Y))) \geqslant E((f(X)g(Y)) \vee (g(X)f(Y)))$	Corollary 5.4, Implied by *SIP*
SAS>	$\sum_{i,j}((a_i + a_j) \vee (b_i + b_j))x_i x_j \geqslant \sum_{i,j}((a_i + b_j) \vee (b_i + a_j))x_i x_j$ $\int\int[(f(x) \vee f(y)) \vee (g(x) \vee g(y))]\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) \geqslant \int\int[(f(x) \vee g(y)) \vee (g(x) \vee f(y))]\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)$ $E((f(X) + f(Y)) \vee (g(X) + g(Y))) \geqslant E((f(X) + g(Y)) \vee (g(X) + f(Y)))$	Implied by *SIS*
AIP<	$\vee_{i,j}((a_i a_j) \wedge (b_i b_j)) \leqslant \vee_{i,j}((a_i b_j) \vee (b_i a_j))$ $\vee_{x,y}((f(x)f(y)) \wedge (g(x)g(y))) \leqslant \vee_{x,y}((f(x)g(y)) \wedge (f(y)g(x)))$	Easy, Section 3
AIS<	$\vee_{i,j}((a_i + a_j) \wedge (b_i + b_j)) \leqslant \vee_{i,j}((a_i + b_j) \vee (b_i + a_j))$ $\vee_{x,y}((f(x) + f(y)) \wedge (g(x) + g(y))) \leqslant \vee_{x,y}((f(x) + g(y)) \wedge (f(y) + g(x)))$	Easy, Section 3
AIA<	$\vee_{i,j}((a_i \vee a_j) \wedge (b_i \vee b_j)) \leqslant \vee_{i,j}((a_i \vee b_j) \wedge (b_i \vee a_j))$ $\vee_{x,y}((f(x) \vee f(y)) \wedge (g(x) \vee g(y))) \leqslant \vee_{x,y}((f(x) \vee g(y)) \wedge (f(y) \vee g(x)))$	Easy, Section 3
API>	$\vee_{i,j}((a_i \wedge a_j)(b_i \wedge b_j)) \geqslant \vee_{i,j}((a_i \wedge b_j)(b_i \wedge a_j))$ $\vee_{x,y}((f(x) \wedge f(y))(g(x) \wedge g(y))) \geqslant \vee_{x,y}((f(x) \wedge g(y))(f(y) \wedge g(x)))$	Easy, Section 3
APS<	$\vee_{i,j}((a_i + a_j)(b_i + b_j)) \leqslant \vee_{i,j}((a_i + b_j)(b_i + a_j))$ $\vee_{x,y}((f(x) + f(y))(g(x) + g(y))) \leqslant \vee_{x,y}((f(x) + g(y))(f(y) + g(x)))$	Easy, Section 3
APA<	$\vee_{i,j}((a_i \vee a_j)(b_i \vee b_j)) \leqslant \vee_{i,j}((a_i \vee b_j)(b_i \vee a_j))$ $\vee_{x,y}((f(x) \vee f(y))(g(x) \vee g(y))) \leqslant \vee_{x,y}((f(x) \vee g(y))(f(y) \vee g(x)))$	Easy, Section 3
ASI>	$\vee_{i,j}((a_i \wedge a_j) + (b_i \wedge b_j)) \geqslant \vee_{i,j}((a_i \wedge b_j) + (b_i \wedge a_j))$ $\vee_{x,y}((f(x) \wedge f(y)) + (g(x) \wedge g(y))) \geqslant \vee_{x,y}((f(x) \wedge g(y)) + (f(y) \wedge g(x)))$	Easy, Section 3
ASP>	$\vee_{i,j}((a_i a_j) + (b_i b_j)) \geqslant \vee_{i,j}((a_i b_j) + (b_i a_j))$ $\vee_{x,y}((f(x)f(y)) + (g(x)g(y))) \geqslant \vee_{x,y}((f(x)g(y)) + (f(y)g(x)))$	Easy, Section 3

Table 1 (*continued*)

DEF	Explicit form and generalizations (when possible)	Proof
ASA<	$\vee_{i,j}((a_i \vee a_j) + (b_i \vee b_j)) \leqslant \vee_{i,j}((a_i \vee b_j) + (b_i \vee a_j))$ $\vee_{x,y}((f(x) \vee f(y)) + (g(x) \vee g(y))) \leqslant \vee_{x,y}((f(x) \vee g(y)) + (f(y) \vee g(x)))$	Easy, Section 3
AAI>	$\vee_{i,j}((a_i \wedge a_j) \vee (b_i \wedge b_j)) \geqslant \vee_{i,j}((a_i \wedge b_j) \vee (b_i \wedge a_j))$ $\vee_{x,y}((f(x) \wedge f(y)) \vee (g(x) \wedge g(y))) \geqslant \vee_{x,y}((f(x) \wedge g(y)) \vee (f(y) \wedge g(x)))$	Easy, Section 3
AAP>	$\vee_{i,j}((a_i a_j) \vee (b_i b_j)) \geqslant \vee_{i,j}((a_i b_j) \vee (b_i a_j))$ $\vee_{x,y}((f(x)f(y)) \vee (g(x)g(y))) \geqslant \vee_{x,y}((f(x)g(y)) \vee (f(y)g(x)))$	Easy, Section 3
AAS>	$\vee_{i,j}((a_i + a_j) \vee (b_i + b_j)) \geqslant \vee_{i,j}((a_i + b_j) \vee (b_i + a_j))$ $\vee_{x,y}((f(x) + f(y)) \vee (g(x) + g(y))) \geqslant \vee_{x,y}((f(x) + g(y)) \vee (f(y) + g(x)))$	Easy, Section 3

Assume that $a > 0$, $b > 0$; in some cases, this condition can be relaxed. Assume f and g are measurable and non-negative (or positive, where positivity is required for the expressions to make sense).

1.3. Results

Table 1 states explicitly 48 of the 64 inequalities that involve only S, P, I, A, along with some generalizations of these, including two inequalities identified as false. Table 1 omits the 16 equalities *DEF* where $E = F$. For each true inequality *SEF* in Table 1, the corresponding inequalities *REF* and *QEF* pertaining to the spectral radius and quadratic form are true for non-negative $a, b \in \Re^n$. When *SEF* holds for all real (not merely non-negative) $a, b \in \Re^n$, then *QEF* also holds for all real (not merely non-negative) $a, b \in \Re^n$. The inequalities *SPA*, *QPA*, and *RPA* are all false in general (Section 5). If $n = 2$, then all 96 inequalities are true.

Our inequalities yield a nice generalization of (1.1).

Theorem 1.1. *Let $*$ be one of the four operations $+, \times, \wedge$ and $\vee$ on $\Re$. Let $a, b \in \Re^n$. Denote by $a * a$ the $n \times n$ matrix $a_{i,j} = a_i * a_j$. Then the matrix $a * a$ is more different from $b * b$ than $a * b$ is from $b * a$. Precisely, if $\|A\| = \sum_{1 \leqslant i,j \leqslant n} |a_{i,j}|$, then*

$$\|a * a - b * b\| \geqslant \|a * b - b * a\|.$$

Proof. We use the identities $|x - y| = 2(x \vee y) - x - y = x + y - 2(x \wedge y)$.

1. If $x * y = x \wedge y$, then

$$|a_i \wedge a_j - b_i \wedge b_j| = a_i \wedge a_j + b_i \wedge b_j - 2(a_i \wedge a_j \wedge b_i \wedge b_j),$$
$$|a_i \wedge b_j - b_i \wedge a_j| = a_i \wedge b_j + b_i \wedge a_j - 2(a_i \wedge b_j \wedge b_i \wedge a_j).$$

Therefore

$$\begin{aligned}
&\sum_{1\leqslant i,j\leqslant n} |a_i \wedge a_j - b_i \wedge b_j| - \sum_{1\leqslant i,j\leqslant n} |a_i \wedge b_j - b_i \wedge a_j| \\
&\quad = \sum_{1\leqslant i,j\leqslant n} (a_i \wedge a_j + b_i \wedge b_j) - \sum_{1\leqslant i,j\leqslant n} (a_i \wedge b_j + b_i \wedge a_j) \\
&\quad \geqslant 0,
\end{aligned}$$

because of the inequality *SSI*>.

2. Zbăganu [7] proved the case $x * y = xy$. In fact, he proved more:

$$\|a * a - b * b\| - \|a * b - b * a\| \geqslant \left(\sum_{i=1}^{n}(|a_i| - |b_i|)\right)^2.$$

3. If $x * y = x + y$, then

$$\begin{aligned}
|a_i + a_j - (b_i + b_j)| &= a_i + a_j + b_i + b_j - 2(a_i + a_j) \wedge (b_i + b_j), \\
|a_i + b_j - (b_i + a_j)| &= a_i + b_j + b_i + a_j - 2(a_i + b_j) \wedge (b_i + a_j).
\end{aligned}$$

Therefore

$$\begin{aligned}
&\sum_{1\leqslant i,j\leqslant n} |a_i + a_j - (b_i + b_j)| - \sum_{1\leqslant i,j\leqslant n} |a_i + b_j - (b_i + a_j)| \\
&\quad = 2\sum_{1\leqslant i,j\leqslant n} (a_i + b_j) \wedge (b_i + a_j) - 2\sum_{1\leqslant i,j\leqslant n} (a_i + a_j) \wedge (b_i + b_j) \\
&\quad \geqslant 0
\end{aligned}$$

because of the inequality *SIS*<.

4. If $x * y = x \vee y$, then

$$\begin{aligned}
|a_i \vee a_j - b_i \vee b_j| &= 2(a_i \vee a_j \vee b_i \vee b_j) - (a_i \vee a_j + b_i \vee b_j), \\
|a_i \vee b_j - b_i \vee a_j| &= 2(a_i \vee b_j \vee b_i \vee a_j) - (a_i \vee b_j + b_i \vee a_j).
\end{aligned}$$

Therefore

$$\begin{aligned}
&\sum_{1\leqslant i,j\leqslant n} |a_i \vee a_j - b_i \vee b_j| - \sum_{1\leqslant i,j\leqslant n} |a_i \vee b_j - b_i \vee a_j| \\
&\quad = \sum_{1\leqslant i,j\leqslant n} (a_i \vee b_j + b_i \vee a_j) - \sum_{1\leqslant i,j\leqslant n} (a_i \vee a_j + b_i \vee b_j) \\
&\quad \geqslant 0
\end{aligned}$$

because of the inequality *SSA*<. □

Remark. Another byproduct of our inequalities was indicated to us by Victor de la Peña (personal communication, 2000). The inequality $SAS>$ may be written in terms of the independent and identically distributed random variables X, Y, the real-valued functions f, g and the expectation operator $\mathbf{E}$ (distinguishable by context from the earlier use of E for an unspecified one of the binary operations S, P, I, A) as

$$\begin{aligned}&\mathbf{E}((f(X)+f(Y))\vee(g(X)+g(Y)))\\&\quad\geqslant \mathbf{E}((f(X)+g(Y))\vee(g(X)+f(Y))).\end{aligned}$$

If we let $f(x)=x$, $g(x)=-x$, then

$$\mathbf{E}((X+Y)\vee(-X-Y))\geqslant \mathbf{E}((X-Y)\vee(-X+Y))$$

or $\mathbf{E}|X+Y|\geqslant \mathbf{E}|X-Y|$. This is a special case of the inequality (2.1) of Buja et al. [2] for independent and identically distributed scalar real-valued random variables with $n=1$ and $p=1$. By standard techniques, one can prove for the Euclidean norm that $\mathbf{E}\|X+Y\|\geqslant \mathbf{E}\|X-Y\|$ for independent and identically distributed n-dimensional real random vectors X and Y, since $\|x\|$ is an integral of $|\langle x,a\rangle|$ where a belongs to the unit sphere.

1.4. Organization of the paper

This paper has eight sections. Following this introductory Section 1, Section 2 establishes fundamental definitions and some general principles. A notion of equivalence among inequalities is established. A Remark following Theorem 2.9 shows that it is sufficient to investigate only three classes of inequalities: those with three-letter codes IE_qE_r, PE_qE_r and SE_qE_r, where E_q and E_r are defined at (2.8), (2.34), (2.35). Section 3 discusses IE_qE_r and the equivalent AE_qE_r and proves all 24 inequalities IEF and AEF with $E, F\in\{I,P,S,A\}$ and $E\neq F$. Section 4 analyzes $E_pE_qE_r$ when $0<p=q\leqslant r$ and $p=q<0<r$. Theorem 4.11 extends some of the results obtained for $E_pE_pE_r$ to quadratic forms: If $r\geqslant 1$, then $QSE_r<$ holds. If $r<0$, then $QSE_r>$ holds. Section 5 deals with SE_qE_r and Section 6 deals with PE_qE_r. Section 7 presents generalizations and counterexamples, and reviews major open problems remaining from this work. Section 8 gives results, repeatedly used, that are derived from a theorem on the number of zeros of sums of exponential functions. References for all sections and acknowledgments follow Section 8.

Although we have completely analyzed a large number of inequalities, three mysteries remain. First, why should so many of these inequalities be true, given that they were conjectured by formal analogy? Second, why should the methods used to prove those conjectures that are true be so extraordinarily diverse? Third, what differentiates the few conjectures that turned out to be false from the overwhelming majority of others that turned out to be true?

2. Notations and general principles

Let $E(x, y)$ and $F(x, y)$ be functions such that the composite function

$$H(x_1, x_2, x_3, x_4) = E(F(x_1, x_2), F(x_3, x_4)) \tag{2.1}$$

is defined for each choice of the *non-negative* numbers x_r $(r = 1, 2, 3, 4)$. We will assume that both functions E and F are symmetric and non-decreasing (in the usual senses). Thus the value of H remains unchanged under the following operations:

(i) Interchange x_1 and x_2.
(ii) Interchange x_3 and x_4.
(iii) Interchange the pairs (x_1, x_2) and (x_3, x_4).

(ii) is a consequence of (i) and (iii).

Let n be a positive integer, $i, j \in \{1, 2, \ldots, n\}$ and let J be an interval of real numbers. Often, J will be a subset of $[0, \infty)$ but sometimes J may include negative numbers. The definition of J will depend on the functions involved in a particular situation. Let $a = (a_1, \ldots, a_n)$ and $b = (b_1, \ldots, b_n)$ be arbitrary vectors in J^n. For each pair $a, b \in J^n$, define

$$\begin{aligned} u_{i,j} &= H(a_i, a_j, b_i, b_j) = E(F(a_i, a_j), F(b_i, b_j)), \\ v_{i,j} &= H(a_i, b_j, b_i, a_j) = E(F(a_i, b_j), F(b_i, a_j)), \\ w_{i,j} &= v_{i,j} - u_{i,j}. \end{aligned} \tag{2.2}$$

Since E, F are symmetric,

$$u_{i,j} = u_{j,i}, \quad v_{i,j} = v_{j,i}, \quad i, j = 1, \ldots, n. \tag{2.3}$$

Let

$$U = (u_{i,j})_{i,j=1,\ldots,n}, \quad V = (v_{i,j})_{i,j=1,\ldots,n}, \quad W = (w_{i,j})_{i,j=1,\ldots,n} \tag{2.4}$$

denote the associated $n \times n$ symmetric matrices. The scalars $u_{i,j}$, $v_{i,j}$ and matrices $U = U(a, b)$ and $V = V(a, b)$ depend on the pair of vectors $a, b \in J^n$.

We aim to determine when, under various additional assumptions, U is "smaller" (or "larger", respectively) than V for every $a, b \in J^n$.

Definitions. Let D be a function which assigns a real number DW to each symmetric $n \times n$ matrix $W = (w_{i,j})$. Then property *DEF*< holds if

$$DU \leqslant DV \quad \text{for all } n \geqslant 1, \quad \text{and all } a, b \in J^n. \tag{2.5}$$

Property *DEF*> holds if

$$DU \geqslant DV \quad \text{for all } n \geqslant 1, \quad \text{and all } a, b \in J^n. \tag{2.6}$$

Property *DEF*= holds if both (2.5) and (2.6) are true, i.e.,

$$DU = DV \quad \text{for all } n \geqslant 1, \quad \text{and all } a, b \in J^n. \tag{2.7}$$

Remark. Usually DW will be some average of all the elements, or of only the diagonal elements, of W, as illustrated in (2.8).

Equality. If $a = b$, then from (2.2) one has $u_{i,j} = v_{i,j}$ for all i, j, hence $U = V$ and $DU = DV$.

Definitions. D is monotone if DU is non-decreasing in each entry $u_{i,j}$ (with $i \leqslant j$, always insisting on the symmetry $u_{j,i} = u_{i,j}$).

The function $U \mapsto DU$ is symmetric if D is a symmetric function of all the n^2 variables $u_{i,j}$, i.e., if $\tau : \{1, 2, \ldots, n\}^2 \to \{1, 2, \ldots, n\}^2$ is a permutation and τU means the matrix $(u_{\tau(i,j)})_{i,j}$, then $DU = D(\tau U)$.

Remark. If D is symmetric and $b = (b_1, \ldots, b_n)$ can be obtained from $a = (a_1, \ldots, a_n)$ by a permutation σ of $\{1, \ldots, n\}$ which is its own inverse, then $DU = DV$.

Equivalently,

$$b_j = a_{\sigma(j)} \text{ for all } j \text{ and } \sigma^{-1} = \sigma \quad \Rightarrow \quad DU = DV.$$

Proof. By assumption $a_j = a_{\sigma\sigma(j)} = b_{\sigma(j)}$ for all j. It follows that $u_{i,\sigma(j)} = E(F(a_i, a_{\sigma(j)}), F(b_i, b_{\sigma(j)})) = E(F(a_i, b_j), F(b_i, a_j)) = v_{i,j}$ for all i and j. Then $DU = DV$ because D is symmetric. $\square$

A simple example is the case that $b_1 = a_2$ and $b_2 = a_1$, while $b_i = a_i$ for $i \geqslant 3$.

Definitions. The monotone functions S, P, I, A, R stand for "sum", "product", "minimum", "maximum", and "spectral radius", respectively. If $W = (w_{i,j})$ is any $n \times n$ symmetric matrix (including but not limited to that defined in (2.4)) with eigenvalues λ_j, and $i, j \in \{1, 2, \ldots, n\}$, then

$$PW = \prod_{i,j} w_{i,j}, \quad SW = \sum_{i,j} w_{i,j}, \quad IW = \min_{i,j} w_{i,j}, \quad AW = \max_{i,j} w_{i,j},$$
$$RW = \max\{|\lambda_j|, 1 \leqslant j \leqslant n\}, \quad E_pW = \left(\sum_{i,j} w_{i,j}^p\right)^{1/p} \quad \text{for } p \neq 0. \tag{2.8}$$

The operator P applies only to matrices W that are non-negative. Thus, *DEF* with $D = P$ is allowed only if both $E(x, y)$ and $F(x, y)$ take only non-negative values, for all $x, y \in J$.

Here W is always symmetric, so the eigenvalues are real. The *spectrum* of W is the ordered set of its eigenvalues, counting the multiplicities. The smallest eigenvalue is denoted by $\sigma_1(W)$, and the greatest one by $\sigma_2(A)$. The *spectral radius* is

$$\rho W = \max(|\sigma_1(W)|, |\sigma_2(W)|). \tag{2.9}$$

If W is semipositive definite, then all the eigenvalues are non-negative and $\rho(W) = \sigma_2(W)$. For any $n \times n$ matrix A, let $A \geqslant 0$ mean that all the entries of A are non-negative. The Perron–Frobenius theorem [5] asserts that

$$A \geqslant 0 \quad \Rightarrow \quad \rho(A) = \sigma_2(A). \tag{2.10}$$

As a consequence, $A \geqslant 0$ implies that $\sigma_1(A) + \sigma_2(A) \geqslant 0$.

The quadratic form $QW : \Re^n \to \Re$ associated with W is defined by

$$QW(x) = (x'Wx)_{x \in \Re^n} = \sum_{1 \leqslant i,j \leqslant n} w_{i,j} x_i x_j. \tag{2.11}$$

We write $U \prec V$ if $V - U$ is semipositive definite, meaning that $QW(x) \geqslant 0$, $\forall x \in \Re^n$. For each W, QW is a function of x, whereas the monotone functions S, P, I, A, R, E_p each yield a single real number.

We shall also use the notation S, P, I, A for the same four binary operations (sum, product, min, max) applied to pairs of reals, instead of to all the elements of a matrix as in (2.8). Thus

$$P(x, y) = xy, \quad S(x, y) = x + y, \tag{2.12}$$

$$I(x, y) = \min(x, y) = x \wedge y, \quad A(x, y) = \max(x, y) = x \vee y. \tag{2.13}$$

Each of P, S, I and A makes $\Re_+ = [0, \infty)$ or $\Re$ into a commutative semigroup and is associative:

$$E(E(x, y), z) = E(x, E(x, y)),$$
$$\text{for all real } x, y, z, \quad \text{for each } E \in \{P, S, I, A\}. \tag{2.14}$$

Examples. We now illustrate the notation defined in (2.5)–(2.7), (2.12) and (2.13). The property *PEF*< holds if and only if, for all $n \geqslant 1$ and all $a, b \in J^n$,

$$\prod_{i=1}^{n}\prod_{j=1}^{n} E(F(a_i, a_j), F(b_i, b_j)) \leqslant \prod_{i=1}^{n}\prod_{j=1}^{n} E(F(a_i, b_j), F(b_i, a_j)). \tag{2.15}$$

Similarly, *SEF*< means that, for all $n \geqslant 1$ and all $a, b \in J^n$,

$$\sum_{i=1}^{n}\sum_{j=1}^{n} E(F(a_i, a_j), F(b_i, b_j)) \leqslant \sum_{i=1}^{n}\sum_{j=1}^{n} E(F(a_i, b_j), F(b_i, a_j)). \tag{2.16}$$

SEF> means that, for all $n \geqslant 1$ and all $a, b \in J^n$,

$$\sum_{i=1}^{n}\sum_{j=1}^{n} E(F(a_i, a_j), F(b_i, b_j)) \geqslant \sum_{i=1}^{n}\sum_{j=1}^{n} E(F(a_i, b_j), F(b_i, a_j)). \tag{2.17}$$

In (2.15), it is understood that $E(x, y) \geqslant 0$. Similarly, the choice $E = P$ requires that $F(x, y) \geqslant 0$. As a more explicit example, property *PIS*< asserts that

$$\prod_{i=1}^{n}\prod_{j=1}^{n}((a_i+a_j)\wedge(b_i+b_j)) \leqslant \prod_{i=1}^{n}\prod_{j=1}^{n}((a_i+b_j)\wedge(b_i+a_j)) \tag{2.18}$$

for all $n \geqslant 1$ and all $a, b \in \Re_+^n$.

Definitions. Property *QEF*< holds if $U \prec V$, or explicitly

$$\sum_{i=1}^{n}\sum_{j=1}^{n} E(F(a_i, a_j), F(b_i, b_j))x_i x_j \leqslant \sum_{i=1}^{n}\sum_{j=1}^{n} E(F(a_i, b_j), F(b_i, a_j))x_i x_j, \tag{2.19}$$

for all $n \geqslant 1$, all $x \in \Re^n$, and all $a, b \in J^n$.

Property *QEF*> holds if $V \prec U$, or explicitly

$$\sum_{i=1}^{n}\sum_{j=1}^{n} E(F(a_i, a_j), F(b_i, b_j))x_i x_j \geqslant \sum_{i=1}^{n}\sum_{j=1}^{n} E(F(a_i, b_j), F(b_i, a_j))x_i x_j \tag{2.20}$$

for all $n \geqslant 1$, all $x \in \Re^n$, and all $a, b \in J^n$.

Property *REF*< holds if the spectral radius of U is not greater than that of V,

$$REF< \iff \rho(U) \leqslant \rho(V), \quad \forall n \geqslant 1, \quad \forall a, b \in J^n, \tag{2.21}$$

$$REF> \iff \rho(U) \geqslant \rho(V), \quad \forall n \geqslant 1, \quad \forall a, b \in J^n. \tag{2.22}$$

It is obvious that *QEF*< $\Rightarrow$ *SEF*< and *QEF*> $\Rightarrow$ *SEF*>. (Put $x_i = 1$, $\forall i$.) The next fact is less obvious.

Theorem 2.1. *If $U \geqslant 0$, $V \geqslant 0$ and $U \prec V$, then $\rho(U) \leqslant \rho(V)$. As a consequence QEF< implies REF<. Moreover, QEF< implies SEF< and QEF> implies SEF> even if U, V are not non-negative.*

Proof. $U \prec V$ means that $\sum_{1\leqslant i,j\leqslant n} u_{i,j}x_i x_j \leqslant \sum_{1\leqslant i,j\leqslant n} v_{i,j}x_i x_j$ or $x'Ux \leqslant x'Vx$, $\forall x \in \Re^n$, hence $\sup\{x'Ux; x \in \Re^n, |x| = 1\} \leqslant \sup\{x'Vx; x \in \Re^n, |x| = 1\}$. But $\sigma_2(U) = \sup\{x'Ux; x \in \Re^n, |x| = 1\}$ (see e.g. [5]). So $\sigma_2(U) \leqslant \sigma_2(V) \Rightarrow \rho(U) \leqslant \rho(V)$ by the Perron–Frobenius theorem. The second claim is trivial. □

2.1. *The relations $E \subset F$*

DEF< and *DEF*> are rarely both true. Often an effective way of determining which of *DEF*< and *DEF*> (if any) might be true is to examine the special case:

$$a_i = x, \quad b_i = y, \quad i = 1, \ldots, n, \tag{2.23}$$

where x and y are arbitrary non-negative constants. Then, from (2.3), all elements of $U = (u_{i,j})$ equal u and all elements of $V = (v_{i,j})$ equal v, where

$$u = E(F(x, x), F(y, y)), \quad v = E(F(x, y), F(y, x)). \tag{2.24}$$

In this special case, for any "reasonable" operator D (such as S, P, I, A; ruling out trivial operators such as $DU \equiv 0$), one has $DU \leqslant DV$ if and only if $u \leqslant v$. Thus $u \leqslant v$ for all x, y is a necessary condition for $DEF<$ which does not depend on D.

Definitions. Let $E \subset F$ denote $E(F(x, x), F(y, y)) \leqslant E(F(x, y), F(y, x))$ for all $x, y \geqslant 0$, and let $E \supset F$ denote $E(F(x, x), F(y, y)) \geqslant E(F(x, y), F(y, x))$ for all $x, y \geqslant 0$.

If $DEF<$ is true, then $E \subset F$ or equivalently

$$u_{i,i} \leqslant v_{i,i}, \quad \forall n \geqslant 1, \ \ \forall i = 1, \ldots, n, \quad \forall a, b \in \Re^n_+. \tag{2.25}$$

Similarly, if $DEF>$ is true, then $E \supset F$ or equivalently

$$u_{i,i} \geqslant v_{i,i}, \quad \forall n \geqslant 1, \ \ \forall i = 1, \ldots, n, \quad \forall a, b \in \Re^n_+. \tag{2.26}$$

Let $T =$ trace (sum of the diagonal elements of a matrix argument). Clearly (2.25) implies $TU \leqslant TV$ while (2.26) implies $TU \geqslant TV$.

In applications, it is usually very easy to check which of $E \subset F$, $E \supset F$ is true (if any). If neither is true, then neither $DEF<$ nor $DEF>$ can be true, regardless of D. If, for instance, $E \subset F$ is true but not $E \supset F$, then $DEF>$ is false while $DEF<$ may or may not be true; and vice versa. If both $E \subset F$, $E \supset F$ hold, or equivalently if

$$E(F(x, x), F(y, y)) = E(F(x, y), F(y, x)) \quad \text{for all } x, y \geqslant 0, \tag{2.27}$$

then $DEF=$ may be true. In any case, (2.27) implies $TEF=$ (where $T =$ trace).

In the special case $E = F$, (2.27) becomes

$$E(E(x, x), E(y, y)) = E(E(x, y), E(y, x)) \quad \text{for all } x, y \geqslant 0, \tag{2.28}$$

Since $E(x, y)$ is always assumed to be symmetric,

$$x * y = E(x, y) \tag{2.29}$$

defines a commutative operation. With $E = F$, (2.27) can be stated as

$$(x * x) * (y * y) = (x * y) * (x * y). \tag{2.30}$$

A sufficient condition for (2.27) is for the operation $*$ to be commutative and associative, for then it follows from (2.2) that

$$u_{i,j} = (a_i * a_j) * (b_i * b_j) = (a_i * b_j) * (b_i * b_j) = v_{i,j} \quad \text{for all } i, j = 1, \ldots, n, \tag{2.31}$$

that is, for all $a, b \in \Re^n_+$, $U = V$ and therefore $DU = DV$. This proves:

Theorem 2.2. *Suppose $x * y = E(x, y)$ is associative, always assuming E is symmetric or equivalently $*$ is commutative. Then* (2.31) *and DEE= are true for any D whatsoever. In particular SEE=, PEE=, IEE=, AEE=, TEE=, QEE= and REE= all hold.*

Corollary. *In the* 16 *cases DEE where D and E are chosen from* $\{S, P, I, A\}$, *DEE= holds.*

An important class of symmetric associative functions E may be constructed as follows. Let H be a (non-empty) sub-semigroup of the additive semigroup $[-\infty, +\infty)$. That is, $x, y \in H$ implies $x + y \in H$. Typically, $H = [c, \infty)$ with $c \geqslant 0$ (usually $c = 0$); or $H = [-\infty, \infty)$. Let G be a subset of $\mathfrak{R}$ of the same cardinality as H and let $\varphi : G \to H$ be a 1:1 function from G onto H. Then

$$x * y = E_\varphi(x, y) = \varphi^{-1}(\varphi(x) + \varphi(y)) \quad \text{for all } x, y \in G \tag{2.32}$$

defines a commutative associative operation on G.

Examples. Choose $G = [0, \infty)$ and $H = [-\infty, \infty)$ and

$$\varphi(x) = \log x, \quad \varphi^{-1}(z) = e^z. \quad \text{Then } x * y = \exp(\log x + \log y) = xy. \tag{2.33}$$

Or choose $G = H = [0, \infty)$ and $\varphi(x) = x^p$ where $p > 0$. Thus $\varphi^{-1}(z) = z^{1/p}$ and

$$x * y = (x^p + y^p)^{1/p} = E_p(x, y) \tag{2.34}$$

by abuse of our notation in (2.8). In particular, $E_1(x, y) = x + y$. Letting $p \to \infty$ leads to

$$x * y = \max(x, y) = E_\infty(x, y),$$

which is not a special case of (2.32). Alternatively, if $G = (0, \infty)$, $H = (0, \infty)$ and $\varphi(x) = x^{-r}$, where $r > 0$, then

$$x * y = \frac{1}{(x^{-r} + y^{-r})^{1/r}} = E_{-r}(x, y) \quad \text{(say)}. \tag{2.35}$$

Letting $r \to \infty$ leads to the limiting case $x * y = \min(x, y) = E_{-\infty}(x, y)$.

We have seen that $E \subset F$ is a necessary condition for *DEF*< and that $E \supset F$ is a necessary condition for *DEF*> (provided $DU > DV$ when $U > V$). We seek conditions such that $E_\varphi \subset E_\psi$, where these functions are defined in (2.32).

Theorem 2.3. *Let φ and ψ be continuous and* 1 : 1 (*and thus strictly monotone*) *functions from G onto H with $G = H = (0, \infty)$ or $G = H = [0, \infty)$.*

(i) *If φ is increasing, then $E_\varphi \subset E_\psi \Leftrightarrow \varphi \circ \psi^{-1}$ is concave and $E_\varphi \supset E_\psi \Leftrightarrow \varphi \circ \psi^{-1}$ is convex.*

(ii) *If φ is decreasing, then $E_\varphi \subset E_\psi \Leftrightarrow \varphi \circ \psi^{-1}$ is convex and $E_\varphi \supset E_\psi \Leftrightarrow \varphi \circ \psi^{-1}$ is concave.*

Proof. Define $\chi = \varphi \circ \psi^{-1}$. Then

$$E_\varphi(E_\psi(x,x), E_\psi(y,y)) = \varphi^{-1}(\chi(2\psi(x)) + \chi(2\psi(y))),$$
$$E_\varphi(E_\psi(x,y), E_\psi(y,x)) = \varphi^{-1}(2\chi(\psi(x) + \psi(y))).$$

Hence $E \subset F$ is equivalent, if $\xi = \psi(x)$, $\eta = \psi(y)$, to

$$\varphi^{-1}(\chi(2\xi) + \chi(2\eta)) \leqslant \varphi^{-1}(2\chi(\xi + \eta)) \quad \text{for all } \xi > 0, \quad \eta > 0. \tag{2.36}$$

If φ is (necessarily strictly) increasing, this in turn is equivalent to

$$\chi(2\xi) + \chi(2\eta) \leqslant (2\chi(\xi + \eta)) \quad \text{for all } \xi > 0, \quad \eta > 0. \tag{2.37}$$

Letting $s = 2\xi$, $t = 2\eta$, an equivalent inequality is

$$\chi\left(\frac{s+t}{2}\right) \geqslant \frac{\chi(s) + \chi(t)}{2} \quad \text{for all } s, t > 0.$$

Because φ, ψ are continuous and thus measurable, $E_\varphi \subset E_\psi$ is true if and only if $\chi = \varphi\psi^{-1}$ is concave.

If φ is strictly decreasing instead, the opposite inequality holds in (2.37). Then $E_\varphi \subset E_\psi$ if and only if $\chi = \phi\psi^{-1}$ is convex.

For $E_\varphi \supset E_\psi$, the function $\chi = \varphi \circ \psi^{-1}$ must be convex or concave, respectively, depending on whether φ is increasing or decreasing, respectively. □

Remark. Equality (2.27) holds if and only if $\varphi \circ \psi^{-1}$ is both convex and concave, thus linear. Moreover, in case (i), if χ is strictly concave, then (2.36) and (2.37) hold with strict inequality when $\xi \neq \eta$. Equivalently, letting $E = E_\varphi$ and $F = E_\psi$, the inequality $E \subset F$ holds with strict inequality when $x \neq y$. The other case (ii) behaves analogously.

Remark. As a reminder, if χ is strictly increasing, then χ is convex (concave) if and only if χ^{-1} is concave (convex). If χ is strictly decreasing, then χ is convex (concave) if and only if χ^{-1} is convex (concave).

Theorem 2.4. *For the associative operators E_φ and E_ψ,*

$$E_\varphi \subset E_\psi \ \textit{ if and only if } \ E_\psi \supset E_\varphi. \tag{2.38}$$

Proof. Suppose first that φ and ψ are both increasing. Thus $\chi = \varphi \circ \psi^{-1}$ is increasing. From Theorem 2.3, case (i), $E_\varphi \subset E_\psi \Leftrightarrow \chi$ is concave $\Leftrightarrow \chi^{-1} = \psi\phi^{-1}$ is convex (since χ is increasing) $\Leftrightarrow E_\psi \supset E_\varphi$ (from (i) with φ and ψ interchanged).

If φ and ψ are both decreasing (so that $\chi = \varphi \circ \psi^{-1}$ is again increasing), then $E_\varphi \subset E_\psi \Leftrightarrow \varphi \circ \psi^{-1}$ is convex $\Leftrightarrow \psi\varphi^{-1}$ is concave $\Leftrightarrow E_\psi \supset E_\varphi$ (from (ii) with φ and ψ interchanged).

Now suppose φ is decreasing and ψ is increasing (so that $\varphi\psi^{-1}$ is decreasing). Then $E_\varphi \subset E_\psi \Leftrightarrow \varphi\psi^{-1}$ is convex $\Leftrightarrow \psi\varphi^{-1}$ is concave $\Leftrightarrow E_\varphi \supset E_\psi$ (from (i) with φ, ψ interchanged).

The same reasoning works if φ is increasing and ψ is decreasing. □

Theorem 2.5. *Transitivity*:

$$\textit{if } E_\varphi \subset E_\psi \ \textit{ and } \ E_\psi \subset E_\theta, \quad \textit{then } E_\varphi \subset E_\theta. \tag{2.39}$$

Proof. Let $f = \varphi \circ \theta^{-1}$, $g = \varphi \circ \psi^{-1}$, $h = \psi \circ \theta^{-1}$. Then $f = (\varphi \circ \psi^{-1})(\psi \circ \theta^{-1}) = g(h)$. According to Theorem 2.3, we want to prove that if φ is increasing (decreasing), then f is concave (convex). Given that $E_\varphi \subset E_\psi$ and $E_\psi \subset E_\theta$, there are four cases depending on whether φ, ψ are increasing or decreasing.

Suppose that φ is increasing.

If ψ is increasing, then $g = \varphi \circ \psi^{-1}$ is increasing and concave and h is concave. This means that $h(px + qy) \geqslant ph(x) + qh(y)$ for any $p, q \geqslant 0$ such that $p + q = 1$. So $f(px + qy) = g(h(px + qy)) \geqslant g(ph(x) + qh(y))$ (since g is increasing) $\geqslant p(gh)(x) + q(gh)(y)$ (since g is concave) $= pf(x) + qf(y)$; that is, f is concave $\Leftrightarrow E_\varphi \subset E_\theta$.

If ψ is decreasing, then $g = \varphi \circ \psi^{-1}$ is decreasing and concave and h is convex. This means that $h(px + qy) \leqslant ph(x) + qh(y)$ for any $p, q \geqslant 0$ such that $p + q = 1$. So $f(px + qy) = g(h(px + qy)) \geqslant g(ph(x) + qh(y))$ (since g is decreasing) $\geqslant p(gh)(x) + q(gh)(y)$ (since g is concave) $= pf(x) + qf(y)$; that is, f is concave $\Leftrightarrow E_\varphi \subset E_\theta$.

Suppose that φ is decreasing.

If ψ is increasing, then $g = \varphi \circ \psi^{-1}$ is decreasing and convex and h is concave. This means that $h(px + qy) \geqslant ph(x) + qh(y)$ for any $p, q \geqslant 0$ such that $p + q = 1$. So $f(px + qy) = g(h(px + qy)) \leqslant g(ph(x) + qh(y))$ (since g is decreasing) $\leqslant p(gh)(x) + q(gh)(y)$ (since g is convex) $= pf(x) + qf(y)$; that is, f is convex $\Leftrightarrow E_\varphi \subset E_\theta$.

If ψ is decreasing, then $g = \varphi \circ \psi^{-1}$ is increasing and convex and h is convex. This means that $h(px + qy) \leqslant ph(x) + qh(y)$ for any $p, q \geqslant 0$ such that $p + q = 1$. So $f(px + qy) = g(h(px + qy)) \leqslant g(ph(x) + qh(y))$ (since g is increasing) $\leqslant p(gh)(x) + q(gh)(y)$ (since g is convex) $= pf(x) + qf(y)$; that is, f is convex $\Leftrightarrow E_\varphi \subset E_\theta$. □

Theorem 2.6. *For* $\alpha, \beta \in \Re/\{0\}$, *if* $E_\alpha = E_\alpha(x, y)$ *on* $[0, \infty)^2$ *is defined by* (2.34), (2.35), *then*

$$E_\alpha \subset E_\beta \ (\textit{equivalently}, E_\beta \supset E_\alpha) \iff \alpha \leqslant \beta. \tag{2.40}$$

Moreover, if $\alpha < \beta$, *then* $E \subset F$ *with* $E = E_\alpha$ *and* $F = E_\beta$ *holds with strict inequality when* $x \neq y$. *Hence, if* $\alpha < \beta$, *then* $E_\beta \subset E_\alpha$ *is false.*

Proof. Apply Theorem 2.3 and the first Remark after it, choosing $\varphi(x) = x^\alpha$, $\psi(x) = x^\beta$ $(x > 0)$. Thus $\varphi \circ \psi^{-1}(y) = y^{\alpha/\beta}$ $(y > 0)$, which is strictly concave if $-1 < \alpha/\beta < 1$, strictly convex if either $\alpha/\beta < -1$ or else $\alpha/\beta > 1$. □

Remark. In the collection $\{E_\alpha; \alpha \in \Re/\{0\}\}$, the operator $P(x, y) = xy$ fits in very nicely in place of the missing operator E_0. More precisely,

$$E_\alpha \subset P \subset E_\beta \iff \alpha < 0 < \beta. \tag{2.41}$$

Proof. Let $\alpha < 0$. The operator E_α coincides with E_φ where $\varphi(x) = x^\alpha$. Further $P = E_\psi$ where $\psi(x) = \log x$, thus $\psi^{-1}(y) = e^y$. Hence $\varphi \circ \psi^{-1}(y) = e^{\alpha y}$ which is convex. Since $\varphi(x) = x^\alpha$ is decreasing, $E_\alpha \subset P$ follows from (ii) of Theorem 2.3.

Let $\beta > 0$ and $\varphi(x) = x^\beta$; thus $E_\beta = E_\varphi$ while $\varphi^{-1}(y) = y^{1/\beta}$. Let $\psi(x) = \log x$; thus $\psi \circ \varphi^{-1}(y) = \log y^{1/\beta} = (1/\beta)\log y$ is concave. Since ψ is increasing, this proves that $P = E_\psi \subset E_\varphi = E_\beta$. □

Thus we should regard P as some sort of limit of E_β as $\beta \to 0$ from below or above, except that the latter limit does not exist in the obvious sense. However, P is a limit of E_ψ for $\beta \to 0$ when we define

$$\psi(x) = \psi_\beta(x) = \frac{x^\beta - 1}{\beta} = \log x + \frac{1}{2}\beta(\log x)^2 + \cdots$$

Then $\psi^{-1}(y) = (1 + \beta y)^{1/\beta}$. Thus

$$\begin{aligned}\psi^{-1}(\psi(x) + \psi(y)) &= \psi^{-1}\left(\frac{x^\beta + y^\beta - 2}{\beta}\right)\\ &= (x^\beta + y^\beta - 1)^{1/\beta} \to xy \quad \text{as } \beta \to 0.\end{aligned}$$

After all, $(1/\beta)\log(x^\beta + y^\beta - 1) = (1/\beta)\log(1 + \beta\log x + \beta\log y + \mathrm{O}(\beta^2)) \to \log xy$.

Remark. The functions $\varphi(x)$ and $\psi(x) = \rho\varphi(x) + \sigma$ (with constants $\rho \neq 0$ and σ) are essentially equivalent in the sense that $E_\varphi = E_\psi$, i.e., both $E_\varphi \subset E_\psi$ and $E_\varphi \supset E_\psi$ (the latter being equivalent to $E_\psi \subset E_\varphi$). After all, $\varphi\psi^{-1}(y) = (y - \sigma)/\rho$ and $\psi \circ \varphi^{-1}(y) = \rho y + \sigma$. The latter two functions, being linear, are both convex and concave.

We thus arrive at the (rough) identification that

$$I = E_{-\infty}, \quad H = E_{-1}, \quad P = E_0, \quad S = E_1, \quad A = E_{+\infty}. \tag{2.42}$$

Here H stands for the harmonic operator

$$H(x, y) = \frac{1}{1/x + 1/y} = \frac{xy}{x + y}. \tag{2.43}$$

It thus follows from (2.33) and (2.35) that (with $E_2(x, y) = \sqrt{x^2 + y^2}$)

$$I \subset H \subset P \subset S \subset E_2 \subset A. \tag{2.44}$$

Our goal is to prove (or disprove) properties of the type *DEF*< or *DEF*> defined in (2.5) and (2.6). In the sequel, we will restrict E and F to associative, commutative operators $E = E_\varphi$ and $F = E_\psi$ except that we will also include $I = E_{-\infty}$, $A = E_{+\infty}$ and $P = E_0$. To see more precisely why I, A and $P = E_0$ are limiting cases, we analyze the conditions (2.5) and (2.6) when $D = E_p$, $E = E_q$, $F = E_r$ with $p, q, r \in \Re\backslash\{0\}$.

Definitions. Given $a, b \in (0, \infty)^n$ we denote DU by $L_{p,q,r}(a, b)$ and DV by $R_{p,q,r}(a, b)$. Thus if $p, q, r \notin \{-\infty, 0, \infty\}$, we have

$$L_{p,q,r}(a, b) = \left(\sum_{1\leqslant i,j\leqslant n} \left((a_i^r + a_j^r)^{\frac{q}{r}} + (b_i^r + b_j^r)^{\frac{q}{r}}\right)^{\frac{p}{q}}\right)^{\frac{1}{p}}, \tag{2.45}$$

$$R_{p,q,r}(a, b) = \left(\sum_{1\leqslant i,j\leqslant n} \left((a_i^r + b_j^r)^{\frac{q}{r}} + (b_i^r + a_j^r)^{\frac{q}{r}}\right)^{\frac{p}{q}}\right)^{\frac{1}{p}}. \tag{2.46}$$

So

$$\begin{aligned} E_pE_qE_r< &\iff L_{p,q,r}(a, b) \leqslant R_{p,q,r}(a, b), \quad \forall n \geqslant 1, \quad \forall a, b \in (0, \infty)^n, \\ E_pE_qE_r> &\iff L_{p,q,r}(a, b) \geqslant R_{p,q,r}(a, b), \quad \forall n \geqslant 1, \quad \forall a, b \in (0, \infty)^n. \end{aligned} \tag{2.47}$$

When one of the indices p, q, r belongs to $\{-\infty, 0, \infty\}$, then (2.45) becomes

$$L_{p,q,\infty}(a, b) = \left(\sum_{1\leqslant i,j\leqslant n} \left((a_i \vee a_j)^q + (b_i \vee b_j)^q\right)^{\frac{p}{q}}\right)^{\frac{1}{p}}, \tag{2.48}$$

$$L_{p,\infty,r}(a, b) = \left(\sum_{1\leqslant i,j\leqslant n} \left((a_i^r + a_j^r)^{\frac{1}{r}} \vee (b_i^r + b_j^r)^{\frac{1}{r}}\right)^{p}\right)^{\frac{1}{p}}, \tag{2.49}$$

$$L_{\infty,q,r}(a, b) = \max_{1\leqslant i,j\leqslant n} \left((a_i^r + a_j^r)^{\frac{q}{r}} + (b_i^r + b_j^r)^{\frac{q}{r}}\right)^{\frac{1}{q}}, \tag{2.50}$$

$$L_{p,q,-\infty}(a, b) = \left(\sum_{1\leqslant i,j\leqslant n} \left((a_i \wedge a_j)^q + (b_i \wedge b_j)^q\right)^{\frac{p}{q}}\right)^{\frac{1}{p}}, \tag{2.51}$$

$$L_{p,-\infty,r}(a,b)=\left(\sum_{1\leqslant i,j\leqslant n}\left((a_i^r+a_j^r)^{\frac{1}{r}}\wedge(b_i^r+b_j^r)^{\frac{1}{r}}\right)^p\right)^{\frac{1}{p}}, \tag{2.52}$$

$$L_{-\infty,q,r}(a,b)=\min_{1\leqslant i,j\leqslant n}\left((a_i^r+a_j^r)^{\frac{q}{r}}+(b_i^r+b_j^r)^{\frac{q}{r}}\right)^{\frac{1}{q}}, \tag{2.53}$$

$$L_{p,q,0}(a,b)=\left(\sum_{1\leqslant i,j\leqslant n}\left((a_ia_j)^q+(b_ib_j)^q\right)^{\frac{p}{q}}\right)^{\frac{1}{p}}, \tag{2.54}$$

$$L_{p,0,r}(a,b)=\left(\sum_{1\leqslant i,j\leqslant n}\left((a_i^r+a_j^r)^{\frac{1}{r}}(b_i^r+b_j^r)^{\frac{1}{r}}\right)^p\right)^{\frac{1}{p}}, \tag{2.55}$$

$$L_{0,q,r}(a,b)=\prod_{1\leqslant i,j\leqslant n}\left((a_i^r+a_j^r)^{\frac{q}{r}}+(b_i^r+b_j^r)^{\frac{q}{r}}\right)^{\frac{1}{q}}. \tag{2.56}$$

Theorem 2.7. *Limiting cases*:

(i) $\lim_{r\to\infty}L_{p,q,r}(a,b)=L_{p,q,\infty}(a,b)$, $\lim_{r\to-\infty}L_{p,q,r}(a,b)=L_{p,q,-\infty}(a,b)$, $-\infty\leqslant p,q\leqslant+\infty$,

(ii) $\lim_{q\to\infty}L_{p,q,r}(a,b)=L_{p,\infty,r}(a,b)$, $\lim_{q\to-\infty}L_{p,q,r}(a,b)=L_{p,-\infty,r}(a,b)$, $-\infty\leqslant p,r\leqslant+\infty$,

(iii) $\lim_{p\to\infty}L_{p,q,r}(a,b)=L_{\infty,q,r}(a,b)$, $\lim_{p\to-\infty}L_{p,q,r}(a,b)=L_{-\infty,q,r}(a,b)$, $-\infty\leqslant q,r\leqslant+\infty$,

(iv) $\lim_{r\to0}\dfrac{L_{p,q,r}(a,b)}{2^{\frac{1}{r}}}=L_{p,q,0}(\sqrt{a},\sqrt{b})$,

(v) $\lim_{q\to0}\dfrac{L_{p,q,r}(a,b)}{2^{\frac{1}{q}}}=L_{p,0,2r}(\sqrt{a},\sqrt{b})$ [*the presence of* $2r$ *on the right is intentional*],

(vi) $\lim_{p\to0}\dfrac{L_{p,q,r}(a,b)}{n^{\frac{2}{p}}}=(L_{0,q,r}(a,b))^{\frac{1}{n^2}}$ [*n is the dimension of a and b as always*].

The same holds if one replaces "L" with "R". Here $\sqrt{a}$, $\sqrt{b}$ *are the vectors with components* $(\sqrt{a_i})_{1\leqslant i\leqslant n}$, $(\sqrt{b_i})_{1\leqslant i\leqslant n}$.

Proof. If $x\in(0,\infty)^m$, then

$$\lim_{p\to\infty}\left(x_1^p+x_2^p+\cdots+x_m^p\right)^{\frac{1}{p}}=\max\{x_i;1\leqslant i\leqslant m\},$$

$$\lim_{p\to-\infty}\left(x_1^p+x_2^p+\cdots+x_m^p\right)^{\frac{1}{p}}=\min\{x_i;1\leqslant i\leqslant m\}$$

and

$$\lim_{p\to 0}\left(\frac{x_1^p+x_2^p+\cdots+x_m^p}{m}\right)^{\frac{1}{p}}=(x_1x_2\cdots x_m)^{1/m}.$$

For instance, we compute the limit (vi):

$$\begin{aligned}\lim_{p\to 0}\frac{L_{p,q,r}(a,b)}{n^{\frac{2}{p}}}&=\lim_{p\to 0}\left(\frac{\sum_{1\leqslant i,j\leqslant n}\left((a_i^r+a_j^r)^{\frac{q}{r}}+(b_i^r+b_j^r)^{\frac{q}{r}}\right)^{\frac{p}{q}}}{n^2}\right)^{\frac{1}{p}}\\&=\left(\prod_{1\leqslant i,j\leqslant n}\left((a_i^r+a_j^r)^{\frac{q}{r}}+(b_i^r+b_j^r)^{\frac{q}{r}}\right)^{\frac{1}{q}}\right)^{\frac{1}{n^2}}\\&=(L_{0,q,r}(a,b))^{\frac{1}{n^2}}.\end{aligned}$$

The limit (v) with q, the middle index, is subtler:

$$\begin{aligned}\lim_{q\to 0}\frac{L_{p,q,r}(a,b)}{2^{\frac{1}{q}}}&=\lim_{q\to 0}\left(\sum_{1\leqslant i,j\leqslant n}\left(\frac{(a_i^r+a_j^r)^{\frac{q}{r}}+(b_i^r+b_j^r)^{\frac{q}{r}}}{2}\right)^{\frac{p}{q}}\right)^{\frac{1}{p}}\\&=\left(\sum_{1\leqslant i,j\leqslant n}\left(\lim_{q\to 0}\frac{(a_i^r+a_j^r)^{\frac{q}{r}}+(b_i^r+b_j^r)^{\frac{q}{r}}}{2}\right)^{\frac{p}{q}}\right)^{\frac{1}{p}}\\&=\left(\sum_{1\leqslant i,j\leqslant n}\left((a_i^r+a_j^r)^{\frac{1}{2r}}\times(b_i^r+b_j^r)^{\frac{1}{2r}}\right)^{p}\right)^{\frac{1}{p}}\\&=L_{p,0,2r}(\sqrt{a},\sqrt{b}).\qquad\square\end{aligned}$$

Remark. The function $(p,q,r)\mapsto L_{p,q,r}(a,b)$ is not continuous at 0, but there are no problems for the limits at $\pm\infty$, i.e., $\lim_{p\to p_0,q\to q_0,r\to r_0}L_{p,q,r}(a,b)=L_{p_0,q_0,r_0}(a,b)$, $\forall p_0,q_0,r_0\in[-\infty,\infty]\backslash\{0\}$.

The next result considerably simplifies our approach.

Theorem 2.8. *Let $t\neq 0$ be arbitrary and p,q,r any real numbers or $\pm\infty$. Then*

$$L_{tp,tq,tr}(a,b)=(L_{p,q,r}(a^t,b^t))^{1/t},\quad R_{tp,tq,tr}(a,b)=(R_{p,q,r}(a^t,b^t))^{1/t},\tag{2.57}$$

where a^t denotes the vector $(a_i^t)_{1\leqslant i\leqslant n}$ and b^t denotes the vector $(b_i^t)_{1\leqslant i\leqslant n}$.

Proof. If p, q, r are real numbers different from 0, then

$$\begin{aligned} L_{tp,tq,tr}(a,b) &= \left(\sum_{1\leqslant i,j\leqslant n}\left((a_i^{tr}+a_j^{tr})^{\frac{tq}{tr}}+(b_i^{tr}+b_j^{tr})^{\frac{tq}{tr}}\right)^{\frac{tp}{tq}}\right)^{\frac{1}{tp}} \\ &= \left(\left(\sum_{1\leqslant i,j\leqslant n}\left(\left((a_i^t)^r+(a_j^t)^r\right)^{\frac{q}{r}}+\left((b_i^t)^r+(b_j^t)^r\right)^{\frac{q}{r}}\right)^{\frac{p}{q}}\right)^{\frac{1}{p}}\right)^{\frac{1}{t}} \\ &= (L_{p,q,r}(a^t,b^t))^{1/t}. \end{aligned}$$

Another six cases when one or two of the indices p, q, r equal 0 raise no problems, but are space consuming. For instance,

$$\begin{aligned} L_{tp,0,tr}(a,b) &= \left(\sum_{1\leqslant i,j\leqslant n}\left((a_i^{tr}+a_j^{tr})^{\frac{tp}{tr}}(b_i^{tr}+b_j^{tr})^{\frac{tp}{tr}}\right)\right)^{\frac{1}{tp}} \\ &= \left(\left(\sum_{1\leqslant i,j\leqslant n}\left(\left((a_i^t)^r+(a_j^t)^r\right)^{\frac{p}{r}}\left((b_i^t)^r+(b_j^t)^r\right)^{\frac{p}{r}}\right)\right)^{\frac{1}{p}}\right)^{\frac{1}{t}} \\ &= (L_{p,0,r}(a^t,b^t))^{1/t} \end{aligned}$$

and

$$\begin{aligned} L_{0,tq,0}(a,b) &= \prod_{1\leqslant i,j\leqslant n}\left((a_ia_j)^{tq}+(b_ib_j)^{tq}\right)^{\frac{1}{tq}} \\ &= \left(\prod_{1\leqslant i,j\leqslant n}\left((a_i^ta_j^t)^q+(b_i^tb_j^t)^q\right)^{\frac{1}{q}}\right)^{\frac{1}{t}} \\ &= (L_{0,q,0}(a^t,b^t))^{1/t}. \qquad \square \end{aligned}$$

2.2. *Equivalences*

Definitions. The inequalities *DEF* (with either $<$ or $>$) and *GHK* (with either $<$ or $>$) are equivalent if and only if the truth of one implies the truth of the other, and vice versa. An equivalence between *DEF* and *GHK* will be written as $DEF \leftrightarrow GHK$.

Equivalent inequalities must either both be false or else both be true. It is possible that $DEF< \Leftrightarrow GHK<$ or that $DEF< \Leftrightarrow GHK>$. In the former case, we say that

DEF and *GHK* are *similar*. In the latter case, where the direction of the inequalities is reversed, we say that *DEF* and *GHK* are *dual*.

Theorem 2.9. *$E_pE_qE_r \leftrightarrow E_{tp}E_{tq}E_{tr}$, whenever $t \neq 0$. If $t > 0$, then $E_pE_qE_r$ and $E_{tp}E_{tq}E_{tr}$ are similar; if $t < 0$, they are dual.*

Proof. Let $t > 0$. If $E_pE_qE_r<$, then $L_{p,q,r}(a, b) \leqslant R_{p,q,r}(a, b)$, $\forall (a, b) \in (0, \infty)^n$. Then $L_{tp,tq,tr}(a, b) = (L_{p,q,r}(a^t, b^t))^{1/t}$ (from (2.57)) $\leqslant (R_{p,q,r}(a^t, b^t))^{1/t} = R_{tp,tq,tr}(a, b)$, hence $E_{tp}E_{tq}E_{tr}<$ holds. If $E_{tp}E_{tq}E_{tr}<$ holds, then replacing t with $1/t$ we see that $E_pE_qE_r<$ holds, too. In the same way, one checks that $E_pE_qE_r> \Leftrightarrow E_{tp}E_{tq}E_{tr}>$.

If $t < 0$, then $E_pE_qE_r<$ implies that $L_{p,q,r}(a^t, b^t) \leqslant R_{p,q,r}(a^t, b^t) \Rightarrow (L_{p,q,r}(a^t, b^t))^{1/t} \geqslant (R_{p,q,r}(a^t, b^t))^{1/t} \Rightarrow L_{tp,tq,tr}(a, b) \geqslant R_{tp,tq,tr}(a, b)$, $\forall n \geqslant 1$, $\forall a, b \in (0, \infty)^n \Leftrightarrow E_{tp}E_{tq}E_{tr}>$. □

Remark. This theorem implies that if $p \notin \{-\infty, 0, \infty\}$, then the inequality $E_pE_qE_r$ is equivalent to $SE_{q/p}E_{r/p}$. Also, any inequality IE_qE_r is equivalent to $AE_{-q}E_{-r}$. Therefore, we have to study only three types of inequalities: IE_qE_r, PE_qE_r and SE_qE_r. Consequently, Section 3 discusses IE_qE_r, Sections 4 and 5 deal with SE_qE_r and Section 6 deals with PE_qE_r.

Definition. The set Ω denotes the set of all $(p, q, r) \in [-\infty, \infty]^3$ such that $E_pE_qE_r<$ or $E_pE_qE_r>$ hold(s).

One goal of this paper is to find properties of this set Ω.

Theorem 2.10. *If $(p, q, r) \in \Omega$ and $t \in \Re$, then $(tp, tq, tr) \in \Omega$. If $(p_m, q_m, r_m) \to (p, q, r) \in ([-\infty, \infty]\backslash\{0\})^3$ and $(p_m, q_m, r_m) \in \Omega$, then $(p, q, r) \in \Omega$.*

Proof. For $t \neq 0$, the first assertion says that $E_pE_qE_r \leftrightarrow E_{tp}E_{tq}E_{tr}$. If $t = 0$, then $E_{tp}E_{tq}E_{tr} = E_0E_0E_0 = PPP$ which holds with PPP=. So $(0, 0, 0) \in \Omega$. The second assertion comes from continuity of $L_{p,q,r}$ and $R_{p,q,r}$ outside 0 and Theorem 2.7. □

Remark. According to the above theorem, one can visualize the set Ω as sets at three levels of p: at level $p = -\infty$, we get the set $\Omega_{-\infty} = \{(q, r); (-\infty, q, r) \in \Omega\}$; at level $p = 0$, we get the set $\Omega_0 = \{(q, r); (0, q, r) \in \Omega\}$; and, finally, at level $p = 1$, we get the set $\Omega_1 = \{(p, q); (1, p, q) \in \Omega\}$.

Other useful ways to establish equivalence are changing the sign and taking the logarithm.

(i) Changing the sign. Consider *DEF* and suppose $D, E, F \in \{I, S, A\}$ (P is excluded) and $E \neq F$ and at least one of E, F belongs to $\{A, I\}$. Now suppose this *DEF* is true in the sense that either (2.5) or (2.6) holds for each choice of $a, b \in \Re^n_+$. Let $e = (1, \ldots, 1) \in \Re^n_+$. It is easily seen from (2.2) that, on replacing a by $a + \lambda e$ and b by $b + \lambda e$, both DU and DV increase by exactly the same amount (depending on the choice of $\lambda \in \Re$). Hence, the difference $DV - DU$ remains unchanged. It follows that if the inequality at hand ($DU \leqslant DV$ or $DU \geqslant DV$) is true for all $a, b \in \Re^n_+$, then it must also be true for all $a, b \in \Re^n$. Consequently, that inequality remains true under a simultaneous replacement of a by $-a$ and b by $-b$. Pulling the minus sign to the front gives a new (but equivalent) inequality, having opposite direction. Its type can be obtained from *DEF* by changing each I into A and each A into I, while leaving any S unchanged. This method proves nine equivalences.

$$\begin{array}{lll} SIS \leftrightarrow SAS & AIS \leftrightarrow IAS & IIS \leftrightarrow AAS \\ SIA \leftrightarrow SAI & AIA \leftrightarrow IAI & IIA \leftrightarrow AAI \\ SSA \leftrightarrow SSI & ASA \leftrightarrow ISI & ISA \leftrightarrow ASI \end{array} \tag{2.58}$$

(ii) Taking the logarithm. Another strategy is to transform S into P by introducing a logarithm. Consider an equality *DEF*, as in (i) with $D, E, F \in \{I, S, A\}$ (none equal to P) and $E \neq F$ and at least one of E, F belongs to $\{A, I\}$. As we saw, if *DEF* is true for all $a, b \in \Re^n_+$, then it must also be true for all $a, b \in \Re^n$. Now replace each old a_i by $\log a_i$, and each old b_j by $\log b_j$ where the new a_i, b_j are strictly positive. Finally, replace DU and DV by $\exp(DU)$ and $\exp(DV)$. This creates a new (but equivalent) inequality having the same direction as the original *DEF*. Its type can be obtained from *DEF* by replacing each S by P. This method proves 14 equivalences.

$$\begin{array}{lll} SSI \leftrightarrow PPI & SSA \leftrightarrow PPA & SAS \leftrightarrow PAP \\ SIS \leftrightarrow PIP & SAI \leftrightarrow PAI & SIA \leftrightarrow PIA \\ ASA \leftrightarrow APA & ASI \leftrightarrow API & ISA \leftrightarrow IPA \\ ISI \leftrightarrow IPI & AIS \leftrightarrow AIP & IAS \leftrightarrow IAP \\ AAS \leftrightarrow AAP & IIS \leftrightarrow IIP & \end{array} \tag{2.59}$$

Together (2.58) and (2.59) are equivalent to eight quadruplets and two pairs.

$$\begin{array}{l} PPI \leftrightarrow PPA \leftrightarrow SSI \leftrightarrow SSA \\ PIP \leftrightarrow PAP \leftrightarrow SIS \leftrightarrow SAS \\ PIA \leftrightarrow PAI \leftrightarrow SIA \leftrightarrow SAI \\ IPI \leftrightarrow APA \leftrightarrow ISI \leftrightarrow ASA \\ IPA \leftrightarrow API \leftrightarrow ISA \leftrightarrow ASI \\ IAP \leftrightarrow AIP \leftrightarrow IAS \leftrightarrow AIS \\ IIP \leftrightarrow AAP \leftrightarrow IIS \leftrightarrow AAS \\ ISI \leftrightarrow ASA \leftrightarrow IPI \leftrightarrow APA \\ IIA \leftrightarrow AAI \\ AIA \leftrightarrow IAI \end{array} \tag{2.60}$$

The equivalences (2.60) involve $32 + 4 = 36$ properties *DEF* and reduce the proof of these 36 properties to the proof of just $8 + 2 = 10$ properties (a gain of 26). In addition, there are $48 - 36 = 12$ properties *DEF* (with $E \neq F$) not yet mentioned so far.

3. Inequalities of the form *IEF* or *AEF*

We now prove all 24 inequalities *IEF* and *AEF* with $E, F \in \{I, P, S, A\}$ and $E \neq F$. These inequalities are consequences of a more general fact.

Theorem 3.1. *Let* $\varphi, \psi : (0, \infty) \to (0, \infty)$ *be* 1:1 *onto continuous functions* (*hence they are monotone*). *Define*, *as in* (2.32),

$$E_\varphi(x, y) = \varphi^{-1}(\varphi(x) + \varphi(y)), \quad E_\psi(x, y) = \psi^{-1}(\psi(x) + \psi(y)) \\ \text{for all } x, y > 0. \tag{3.1}$$

Then

$$E_\varphi \subset E_\psi \iff IE_\varphi E_\psi < \text{ and } AE_\varphi E_\psi < \text{ hold},$$
$$E_\varphi \supset E_\psi \iff IE_\varphi E_\psi > \text{ and } AE_\varphi E_\psi > \text{ hold}.$$

Proof. For any $n \geqslant 1$, let $a, b \in (0, \infty)^n$, $u_{i,j} = E_\varphi(E_\psi(a_i, a_j), E_\psi(b_i, b_j))$, $v_{i,j} = E_\varphi(E_\psi(a_i, b_j), E_\psi(b_i, a_j))$, $1 \leqslant i, j \leqslant n$. Expanding these expressions gives

$$\begin{aligned} u_{i,j} &= \varphi^{-1}(\varphi(E_\psi(a_i, a_j)) + \varphi(E_\psi(b_i, b_j))) \\ &= \varphi^{-1}(\varphi(\psi^{-1}(\psi(a_i) + \psi(a_j))) + \varphi(\psi^{-1}(\psi(b_i) + \psi(b_j)))), \\ v_{i,j} &= \varphi^{-1}(\varphi(E_\psi(a_i, b_j)) + \varphi(E_\psi(b_i, a_j))) \\ &= \varphi^{-1}(\varphi(\psi^{-1}(\psi(a_i) + \psi(b_j))) + \varphi(\psi^{-1}(\psi(b_i) + \psi(a_j)))). \end{aligned}$$

Denote the mapping $\varphi \circ \psi^{-1}$ by χ, as in the previous section. Let also $x_i = \psi(a_i)$ and $y_i = \psi(b_i)$. Then

$$\begin{aligned} u_{i,j} &= \varphi^{-1}(\chi(x_i + x_j) + \chi(y_i + y_j)), \\ v_{i,j} &= \varphi^{-1}(\chi(x_i + y_j) + \chi(y_i + x_j)). \end{aligned} \tag{3.2}$$

Case 1. Suppose that $E_\varphi \subset E_\psi$. Then $u_{i,i} \leqslant v_{i,i}$, $\forall 1 \leqslant i \leqslant n$. The task is to prove that

$$\bigwedge_{1\leqslant i,j\leqslant n} u_{i,j} \leqslant \bigwedge_{1\leqslant i,j\leqslant n} v_{i,j} \quad \text{and} \quad \bigvee_{1\leqslant i,j\leqslant n} u_{i,j} \leqslant \bigvee_{1\leqslant i,j\leqslant n} v_{i,j}. \tag{3.3}$$

We shall prove more, namely,

$$\bigwedge_{1\leqslant i\leqslant n} u_{i,i} \leqslant \bigwedge_{1\leqslant i,j\leqslant n} v_{i,j} \quad \text{and} \quad \bigvee_{1\leqslant i,j\leqslant n} u_{i,j} \leqslant \bigvee_{1\leqslant i\leqslant n} v_{i,i}. \tag{3.4}$$

To prove (3.4), we shall prove, for any two different indices i, j:

$$u_{i,i} \wedge u_{j,j} \leqslant v_{i,j} \quad \text{and} \quad u_{i,j} \leqslant v_{i,i} \vee v_{j,j}. \tag{3.5}$$

It is obvious that (3.5) together with $u_{i,i} \leqslant v_{i,i}$, $\forall 1 \leqslant i \leqslant n$ imply (3.4), which in turn implies (3.3). We shall apply Theorem 2.3. There are two subcases:

Case 1.1. φ is increasing. Then φ^{-1} is also increasing and (from Theorem 2.3) χ is concave. By (3.2),

$$u_{i,i} = \varphi^{-1}(\chi(2x_i) + \chi(2y_i)) \quad \text{and} \quad u_{j,j} = \varphi^{-1}(\chi(2x_j) + \chi(2y_j)).$$

As

$$\begin{aligned} u_{i,i} \wedge u_{j,j} &= (\varphi^{-1}(\chi(2x_i) + \chi(2y_i))) \wedge (\varphi^{-1}(\chi(2x_j) + \chi(2y_j))) \\ &= \varphi^{-1}((\chi(2x_i) + \chi(2y_i)) \wedge (\chi(2x_j) + \chi(2y_j))) \end{aligned}$$

(since φ^{-1} is increasing), the first of the inequalities (3.5) becomes

$$\begin{aligned} &\varphi^{-1}((\chi(2x_i) + \chi(2y_i)) \wedge (\chi(2x_j) + \chi(2y_j))) \\ &\quad \leqslant \varphi^{-1}(\chi(x_i + y_j) + \chi(y_i + x_j)). \end{aligned} \tag{3.6}$$

Similarly, since φ^{-1} is increasing,

$$\begin{aligned} v_{i,i} \vee v_{j,j} &= \varphi^{-1}(2\chi(x_i + y_i)) \vee \varphi^{-1}(2\chi(x_j + y_j)) \\ &= \varphi^{-1}(2\chi(x_i + y_i) \vee 2\chi(x_j + y_j)), \end{aligned}$$

the second of the inequalities (3.5) becomes

$$\varphi^{-1}(\chi(x_i + x_j) + \chi(y_i + y_j)) \leqslant \varphi^{-1}(2\chi(x_i + y_i) \vee 2\chi(x_j + y_j)). \tag{3.7}$$

Again because φ^{-1} is increasing, (3.6) and (3.7) become

$$(\chi(2x_i) + \chi(2y_i)) \wedge (\chi(2x_j) + \chi(2y_j)) \leqslant \chi(x_i + y_j) + \chi(y_i + x_j), \tag{3.8}$$
$$\chi(x_i + x_j) + \chi(y_i + y_j) \leqslant 2(\chi(x_i + y_i) \vee \chi(x_j + y_j)). \tag{3.9}$$

But as χ is concave,

$$\begin{aligned} &(\chi(2x_i) + \chi(2y_i)) \wedge (\chi(2x_j) + \chi(2y_j)) \\ &\quad \leqslant (\chi(2x_i) + \chi(2y_i) + \chi(2x_j) + \chi(2y_j))/2 \\ &\quad = (\chi(2x_i) + \chi(2y_j))/2 + (\chi(2x_j) + \chi(2y_i))/2 \\ &\quad \leqslant \chi\left(\frac{2x_i + 2y_j}{2}\right) + \chi\left(\frac{2x_j + 2y_i}{2}\right) \\ &\quad = \chi(x_i + y_j) + \chi(y_i + x_j), \end{aligned}$$

hence (3.8) holds. Moreover,

$$\begin{aligned} &(\chi(x_i + x_j) + \chi(y_i + y_j))/2 \\ &\quad \leqslant \chi\left(\frac{x_i + x_j + y_i + y_j}{2}\right) \quad \text{(by the concavity of } \chi) \\ &\quad = \chi\left(\frac{(x_i + y_i) + (x_j + y_j)}{2}\right) \leqslant (\chi(x_i + y_i) \vee \chi(x_j + y_j)) \end{aligned}$$

since χ is monotone increasing if ψ is increasing, and monotone decreasing if ψ is decreasing; and if χ is monotone, $\chi((s+t)/2) \leqslant \chi(s) \vee \chi(t)$. Thus (3.9) holds, too. Therefore (3.5) holds. We have proved $IE_\varphi E_\psi <$ and $AE_\varphi E_\psi <$ in this case.

Case 1.2. φ is decreasing. Then φ^{-1} is increasing as well and (from Theorem 2.3) χ is convex. As

$$\begin{aligned} u_{i,i} \wedge u_{j,j} &= (\varphi^{-1}(\chi(2x_i) + \chi(2y_i))) \wedge (\varphi^{-1}(\chi(2x_j) + \chi(2y_j))) \\ &= \varphi^{-1}((\chi(2x_i) + \chi(2y_i)) \vee (\chi(2x_j) + \chi(2y_j))) \end{aligned}$$

(since φ^{-1} is decreasing), the first of the inequalities (3.5) becomes

$$\begin{aligned} &\varphi^{-1}((\chi(2x_i) + \chi(2y_i)) \vee (\chi(2x_j) + \chi(2y_j))) \\ &\quad \leqslant \varphi^{-1}(\chi(x_i + y_j) + \chi(y_i + x_j)). \end{aligned} \tag{3.10}$$

Similarly, since φ^{-1} is decreasing,

$$\begin{aligned} v_{i,i} \vee v_{j,j} &= \varphi^{-1}(2\chi(x_i + y_i)) \vee \varphi^{-1}(2\chi(x_j + y_j)) \\ &= \varphi^{-1}(2\chi(x_i + y_i) \wedge 2\chi(x_j + y_j)), \end{aligned}$$

so the second of the inequalities (3.5) becomes

$$\varphi^{-1}(\chi(x_i + x_j) + \chi(y_i + y_j)) \leqslant \varphi^{-1}(2\chi(x_i + y_i) \wedge 2\chi(x_j + y_j)). \tag{3.11}$$

As φ^{-1} is decreasing, the inequalities (3.10) and (3.11) are equivalent to

$$(\chi(2x_i) + \chi(2y_i)) \vee (\chi(2x_j) + \chi(2y_j)) \geqslant \chi(x_i + y_j) + \chi(y_i + x_j), \tag{3.12}$$

$$\chi(x_i + x_j) + \chi(y_i + y_j) \geqslant 2\chi(x_i + y_i)) \wedge 2\chi(x_j + y_j)). \tag{3.13}$$

The first one (3.12) follows from the convexity of χ:

$$\begin{aligned} &(\chi(2x_i) + \chi(2y_i)) \vee (\chi(2x_j) + \chi(2y_j)) \\ &\quad \geqslant (\chi(2x_i) + \chi(2y_i) + \chi(2x_j) + \chi(2y_j))/2 \\ &\quad = (\chi(2x_i) + \chi(2y_j))/2 + (\chi(2x_j) + \chi(2y_i))/2 \\ &\quad \geqslant \chi(x_i + y_j) + \chi(y_i + x_j). \end{aligned}$$

The second one (3.13) follows from the convexity *and monotonicity* of χ:

$$\begin{aligned} (\chi(x_i + x_j) + \chi(y_i + y_j))/2 &\geqslant \chi\left(\frac{x_i + x_j + y_i + y_j}{2}\right) \\ &= \chi\left(\frac{(x_i + y_i) + (x_j + y_j)}{2}\right) \\ &\geqslant \chi(x_i + y_i) \wedge \chi(x_j + y_j). \end{aligned}$$

Thus we have proved $IE_\varphi E_\psi <$ and $AE_\varphi E_\psi <$ in this case, too.

Case 2. Suppose that $E_\varphi \supset E_\psi$. Then $u_{i,i} \geqslant v_{i,i}$, $\forall 1 \leqslant i \leqslant n$. To prove that

$$\bigwedge_{1\leqslant i,j\leqslant n} u_{i,j} \geqslant \bigwedge_{1\leqslant i,j\leqslant n} v_{i,j} \quad \text{and} \quad \bigvee_{1\leqslant i,j\leqslant n} u_{i,j} \geqslant \bigvee_{1\leqslant i,j\leqslant n} v_{i,j} \tag{3.14}$$

we shall show that for any two different indices i, j we have:

$$u_{i,j} \geqslant v_{i,i} \wedge v_{j,j} \quad \text{and} \quad u_{i,i} \vee u_{j,j} \geqslant v_{i,j}. \tag{3.15}$$

Case 2.1. φ is increasing. From Theorem 2.3, we know that χ is convex. It is also monotone. The proof is similar to that of case 1.1, but one must switch the sense of the inequalities. Then (3.15) is equivalent to

$$\chi(x_i + x_j) + \chi(y_i + y_j) \geqslant 2\chi(x_i + y_j) \wedge 2\chi(x_j + y_j), \tag{3.16}$$

$$(\chi(2x_i) + \chi(2y_i)) \vee (\chi(2x_j) + \chi(2y_j)) \geqslant \chi(x_i + y_j) + \chi(y_i + x_j). \tag{3.17}$$

But (3.16) is a consequence of the convexity and monotonicity of χ:

$$\begin{aligned}(\chi(x_i + x_j) + \chi(y_i + y_j))/2 &\geqslant \chi\left(\frac{x_i + x_j + y_i + y_j}{2}\right)\\ &= \chi\left(\frac{(x_i + y_i) + (x_j + y_j)}{2}\right)\\ &\geqslant \chi(x_i + y_i) \wedge \chi(x_j + y_j).\end{aligned}$$

Likewise, (3.17) is a consequence of the convexity of χ:

$$\begin{aligned}&(\chi(2x_i) + \chi(2y_i)) \vee (\chi(2x_j) + \chi(2y_j))\\ &\quad\geqslant (\chi(2x_i) + \chi(2y_i) + \chi(2x_j) + \chi(2y_j))/2\\ &\quad= (\chi(2x_i) + \chi(2y_i))/2 + (\chi(2x_j) + \chi(2y_i))/2\\ &\quad\geqslant \chi(x_i + y_j) + \chi(y_i + x_j).\end{aligned}$$

So we proved $IE_\varphi E_\psi <$ and $AE_\varphi E_\psi <$ in this case.

Case 2.2. φ is decreasing. From Theorem 2.3, we know that χ is concave and monotone. The proof is similar to that of case 1.2 and (3.15) is equivalent to

$$\chi(x_i + x_j) + \chi(y_i + y_j) \leqslant 2\chi(x_i + y_i) \vee 2\chi(x_j + y_j), \tag{3.18}$$

$$(\chi(2x_i) + \chi(2y_i)) \wedge (\chi(2x_j) + \chi(2y_j)) \leqslant \chi(x_i + y_j) + \chi(y_i + x_j). \tag{3.19}$$

By the concavity and monotonicity of χ,

$$\begin{aligned}(\chi(x_i + x_j) + \chi(y_i + y_j))/2 &\leqslant \chi\left(\frac{x_i + x_j + y_i + y_j}{2}\right)\\ &= \chi\left(\frac{(x_i + y_i) + (x_j + y_j)}{2}\right)\\ &\leqslant \chi(x_i + y_i) \vee \chi(x_j + y_j)\end{aligned}$$

and by concavity,

$$\begin{aligned}
&(\chi(2x_i) + \chi(2y_i)) \wedge (\chi(2x_j) + \chi(2y_j)) \\
&\quad \leqslant (\chi(2x_i) + \chi(2y_i) + \chi(2x_j) + \chi(2y_j))/2 \\
&\quad \leqslant \chi(x_i + y_j) + \chi(y_i + x_j). \qquad \square
\end{aligned}$$

Corollary 3.2. *Denote by* E_p *the operation* E_φ *with* $\varphi(x) = x^p$, $p \in \Re\backslash\{0\}$ (*as in Theorem* 2.6). *From the previous section*, $E_0(s, t) = st$. *Then*

$$p < q \quad \Rightarrow \quad IE_pE_q< \text{ and } AE_pE_q< \text{ hold}, \tag{3.20}$$

$$p > q \quad \Rightarrow \quad IE_pE_q> \text{ and } AE_pE_q> \text{ hold}. \tag{3.21}$$

Proof. According to Theorem 2.6, $p \leqslant q \Leftrightarrow E_p \subset E_q$ and $p \geqslant q \Leftrightarrow E_p \supset E_q$. $\square$

Corollary 3.3. *The following* 24 *inequalities hold*: $IIP<$, $AIP<$, $IIS<$, $AIS<$, $IIA<$, $AIA<$, $IPS<$, $APS<$, $IPA<$, $APA<$, $ISA<$, $ASA<$, $IPI>$, $API>$, $ISI>$, $ASI>$, $IAI>$, $AAI>$, $ISP>$, $ASP>$, $IAP>$, $AAP>$, $IAS>$, $AAS>$.

Proof. According to the previous section, $I = E_{-\infty}$, $P = E_0$, $S = E_1$ and $A = E_\infty$. The conclusion follows from the previous corollary. $\square$

4. Inequalities of the form $E_pE_qE_r$ with $p \leqslant q \leqslant r$ or $p \geqslant q \geqslant r$

If $p \leqslant q \leqslant r$ or $p \geqslant q \geqslant r$, the inequality $E_pE_qE_r<$ holds for $n = 1$ as a consequence of Theorem 2.6. Sometimes the inequality holds more generally. In this section, we consider two cases, $0 < p \leqslant q \leqslant r$ (with results only in the special case $p = q$) and $p = q < 0 < r$. Theorem 4.11 extends some of the results obtained for $E_pE_qE_r<$ to quadratic forms: If $r \geqslant 1$, then $QSE_r<$ holds. If $r < 0$, then $QSE_r>$ holds.

Monotonicity conjecture. *We believe that if* $0 < p \leqslant q \leqslant r$, *then* $E_pE_qE_r<$ *holds.*

We shall prove this conjecture (in Theorem 4.6) when $p = q$, but for the moment assume only $p \leqslant q$. Let $q = \beta r$, $0 \leqslant \beta \leqslant 1$. Denote a_i^r by x_i and b_i^r by y_i. Then

$$u_{i,j} = \left((x_i + x_j)^\beta + (y_i + y_j)^\beta\right)^{\frac{1}{q}}, \quad v_{i,j} = \left((x_i + y_j)^\beta + (y_i + x_j)^\beta\right)^{\frac{1}{q}}. \tag{4.1}$$

Then $E_pE_qE_r<$ means that for all $n \geqslant 1$,

$$\left(\sum_{1 \leqslant i,j \leqslant n} u_{i,j}{}^p\right)^{\frac{1}{p}} \leqslant \left(\sum_{1 \leqslant i,j \leqslant n} v_{i,j}{}^p\right)^{\frac{1}{p}}. \tag{4.2}$$

If we denote $p = \alpha q$, then we must prove that

$$\sum_{1\leqslant i,j\leqslant n} \big((x_i + x_j)^\beta + (y_i + y_j)^\beta\big)^\alpha \leqslant \sum_{1\leqslant i,j\leqslant n} \big((x_i + y_j)^\beta + (y_i + x_j)^\beta\big)^\alpha \tag{4.3}$$

for all $0 \leqslant \alpha, \beta \leqslant 1$. If $\beta = 1$, (4.3) is obvious (and we already knew from Section 2 that $DEE{=}$ always holds).

According to Theorem 2.9, $E_pE_qE_r{<}$ is similar to $SE_{q/p}E_{r/p} = SE_{1/\alpha}E_{1/\alpha\beta}$. We shall prove (4.3) when $p = q$ ($\alpha = 1$). Then the task is to prove that

$$\sum_{1\leqslant i,j\leqslant n} \big((x_i + x_j)^\beta + (y_i + y_j)^\beta\big) \leqslant \sum_{1\leqslant i,j\leqslant n} \big((x_i + y_j)^\beta + (y_i + x_j)^\beta\big), \quad \forall 0 \leqslant \beta \leqslant 1, \tag{4.4}$$

which is $SSE_{1/\beta}{<}$ after replacing x_i by $a_i^{1/\beta}$ and y_i by $b_i^{1/\beta}$. The difference between the left side and the right side of (4.4),

$$D(x, y) = \sum_{1\leqslant i,j\leqslant n} \big((x_i + x_j)^\beta + (y_i + y_j)^\beta - (x_i + y_j)^\beta - (y_i + x_j)^\beta\big), \tag{4.5}$$

has the form

$$D(x, y) = \sum_{1\leqslant i,j\leqslant n} \big(\varphi(x_i, x_j) + \varphi(y_i, y_j) - \varphi(x_i, y_j) - \varphi(y_i, x_j)\big) \tag{4.6}$$

with

$$\varphi(s, t) = f(s + t), \quad f(u) = u^\beta, \ \ 0 \leqslant \beta \leqslant 1. \tag{4.7}$$

This observation motivates the following:

Definitions. Let $D_f : (0,\infty)^n \times (0,\infty)^n \to \Re$ be defined by

$$D_f(x, y) = \sum_{1\leqslant i,j\leqslant n} \big(f(x_i + x_j) + f(y_i + y_j) - f(x_i + y_j) - f(y_i + x_j)\big). \tag{4.8}$$

Let C_1 be the set of all functions $f : (0,\infty) \to \Re$ such that $D_f \geqslant 0$.

According to Proposition 8.3, the set C_1 is a cone with property (A) defined in Section 8. Thus the functions $f(u) = u^\beta$ belong to C_1.

Remark on the case of equality. The functions $f_t(x) = e^{tx}$ belong to C_1 even if $t > 0$. If $D_{f_t}(x, y) = 0$ for any $t \in \Re$, then y is a permutation of x; and conversely.

Proof. Suppose that x_k is the maximum of x_i and that y_m is the maximum of y_j. We have $D_{f_t}(x, y) = 0$ for any $t \in \Re \Leftrightarrow \sum_{i=1}^n e^{tx_i} = \sum_{i=1}^n e^{ty_i}$ for all t. Dividing both sides by e^{tx_k} and taking $t \to \infty$, the left side is a positive number but the right side converges either to 0 (if $x_k > y_m$) or to ∞ (if $x_k < y_m$). As both sides must remain equal, the only possibility is that $x_k = y_m$. Then these terms cancel each other and we get two sums with $n - 1$ terms and repeat the same reasoning. The converse is obvious. $\square$

Theorem 4.1. *If* $0 \leqslant p \leqslant r$, *then* $E_pE_pE_r<$ *holds.*

Proof. If $p > 0$, property $E_pE_pE_r<$ means that $u \to u^\beta$ belongs to C_1, which is assured by Proposition 8.3.

As C_1 has the property (A), it contains the functions $x \to -\log x$, according to Corollary 8.2. Now if $p = 0$, the property $E_pE_pE_r<$ becomes $PPE_r<$, which is equivalent with PPS by Theorem 2.9. But $PPS<$ means that

$$\prod_{1\leqslant i,j\leqslant n} (a_i + a_j)(b_i + b_j) \leqslant \prod_{1\leqslant i,j\leqslant n} (a_i + b_j)(b_i + a_j) \tag{4.9}$$

or, after taking the logarithm,

$$\sum_{1\leqslant i,j\leqslant n} \big(\log(a_i + a_j) + \log(b_i + b_j) - \log(a_i + b_j) - \log(b_i + a_j)\big) \leqslant 0. \tag{4.10}$$

But (4.10) is true because $x \to -\log x$ belongs to C_1. $\square$

Theorem 4.2. *If* $p \leqslant 0 \leqslant r$, *then* $E_pE_pE_r<$ *holds.*

Proof. If $p = 0$, Theorem 4.1 applies.

Let $p < 0 < r$. Since $x \to x^p$ is decreasing, (4.2) is equivalent to

$$\sum_{1\leqslant i,j\leqslant n} {u_{i,j}}^p \geqslant \sum_{1\leqslant i,j\leqslant n} {v_{i,j}}^p \tag{4.11}$$

or, from (4.1) with $\beta = q/r = p/r < 0$, to

$$\sum_{1\leqslant i,j\leqslant n} \big((x_i + x_j)^\beta + (y_i + y_j)^\beta\big) \geqslant \sum_{1\leqslant i,j\leqslant n} \big((x_i + y_j)^\beta + (y_i + x_j)^\beta\big), \tag{4.12}$$

meaning that $x \to -x^\beta$ is in C_1 which is true according to Corollary 8.2.

If $r = 0$, apply Theorem 2.7(iv), passing to limit as $r \downarrow 0$. If $L_{p,p,r}(a, b) \leqslant R_{p,p,r}(a, b)$ for any $r > 0$ and $a, b \geqslant 0$, then

$$\begin{aligned}
&\lim_{r\to 0} \frac{L_{p,p,r}}{2^{\frac{1}{r}}}(a, b) \leqslant \lim_{r\to 0} \frac{R_{p,p,r}}{2^{\frac{1}{r}}}(a, b), \quad \forall a, b \in [0, \infty)^n \\
&\iff L_{p,p,0}(\sqrt{a}, \sqrt{b}) \leqslant R_{p,p,0}(\sqrt{a}, \sqrt{b}), \quad \forall a, b \in [0, \infty)^n \\
&\iff E_pE_pE_0\!< \text{ holds,}
\end{aligned}$$

since any $a, b \geqslant 0$ can be written in the form $(\sqrt{a}, \sqrt{b})$. □

Theorem 4.3

(i) *If* $0 < r < 1/2$, *then* $E_1E_1E_r\!>$ (*equivalently*, $SSE_r\!>$) *is false. However*, *if* $r = 1/2$, *then* $E_1E_1E_r\!>$ *is true.*

(ii) *Let* $p \leqslant r$. *The inequality* $E_pE_pE_r\!<$ *holds if* $r \geqslant 0$. *If* $r < 0$, *the inequality may be false. If* $p \geqslant r$, *then* $E_pE_pE_r\!>$ *holds if* $p \leqslant 0$. *If* $p \geqslant r > 0$, *the inequality may be false.*

(iii) *If* $r \geqslant 1$, $SSE_r\!<$ *holds. If* $r \leqslant 0$, $SSE_r\!>$ *holds.*

(iv) *The set* Ω *of all* $(p, q, r) \in [-\infty, \infty]^3$ *such that* $E_pE_qE_r\!<$ *or* $E_pE_qE_r\!>$ *hold(s)* (*defined before Theorem* 2.10) *contains the set* $\{\pm(p, p, r);\ p \leqslant r$ *and* $r \geqslant 0\}$ *and does not contain the points* $(\beta, \beta, 1)$ *or* $(1, 1, 1/\beta)$ *if* $\beta > 2$. *In particular,* Ω *contains all* $(1, 1, r)$ *with* $r \in [-\infty, 0] \cup [1, \infty] \cup \{1/2\}$.

Proof. (i) Let $r = 1/\beta$. We prove that $E_1E_1E_r\!>$ is false by showing that we can produce a pair $a, b \in [0, \infty)^n$ such that $D < 0$ where

$$\begin{aligned}
D &= D(x, y, n, \beta) \\
&:= \sum_{1\leqslant i,j\leqslant n} \left((x_i + x_j)^\beta + (y_i + y_j)^\beta\right) - \sum_{1\leqslant i,j\leqslant n} \left((x_i + y_j)^\beta + (y_i + x_j)^\beta\right).
\end{aligned} \tag{4.13}$$

Here $x_i = a_i^r$ and $y_i = b_i^r$. We choose $x = (1/2, 1/2, \ldots, 1/2)$ and $y = (1, 0, \ldots, 0)$. Then (4.13) becomes

$$D = n^2 + 2^\beta + (2n - 2) - 2n\left[\left(\frac{3}{2}\right)^\beta + \frac{n-1}{2^\beta}\right]. \tag{4.14}$$

Define

$$g_n(\beta) = 2^\beta D = 4^\beta - 2n \cdot 3^\beta + (n^2 + 2n - 2) \cdot 2^\beta - 2n(n-1) \cdot 1^\beta. \tag{4.15}$$

This way of writing g_n shows that g_n is a particular case of g from Lemma 8.1 with $m = 4$ and $a_1 = \ln 4$, $a_2 = \ln 3$, $a_3 = \ln 2$ and $a_4 = \ln 1 = 0$. According to Lemma 8.1, $g_n = 0$ has at most 3 solutions. The values of g_n which interest us are:

β	$g_n(\beta)$
0	$-(n-1)^2$
1	0
2	$2(n-2)^2$
3	$6[(n-3)^2-1]$
4	$2(7n^2-64n+112)$
5	$30(n^2-14n+32)$
6	$62n^2-1328n+3968$

From this table, we see that $g_2(1) = g_2(2) = g_2(3) = 0$. Therefore $g_2 = 0$ has only the solutions $x_1 = 1, x_2 = 2, x_3 = 3$; so g_2 does not change sign on the interval $(2, 3)$. But $g_2(2.5) < 0$, hence $g_2(\beta) < 0$, $\forall 2 < \beta < 3$ or, equivalently, if $r \in (1/3, 1/2)$.

Let now $n = 3$. From our table, we see that $x_1 = 1$ is the first root of g_3. As $g_3(2) = 2 > 0$ and $g_3(3) = -6 < 0$, another root x_2 is in the interval $(2, 3)$. Also, as $g_3(5) = -30 < 0$, $g_3(6) = 746 > 0$, a root x_3 is in $(5, 6)$. Moreover, between 3 and 5, g_3 does not change sign. Therefore if $\beta \in [3, 5]$ or $r \in [1/5, 1/3]$, $D < 0$.

Finally, let $\beta > 5$. Now write g_n as

$$h_\beta(n) = n^2(2^\beta - 2) - 2n(3^\beta - 2^\beta - 1) + 2^\beta(2^\beta - 2). \tag{4.16}$$

We want to prove that for any $\beta > 5$, there exists an $n = n(\beta)$ such that $h_\beta(n) < 0$, that is, that there exists a positive integer $n \geqslant 2$ such that $n_1 < n < n_2$ where n_1 and n_2 are the two roots of the equation of second degree $h_\beta = 0$. If we divide (4.16) by $2^\beta - 2$, the equation $h_\beta(n) = 0$ becomes

$$n^2 - 2n\frac{3^\beta - 2^\beta - 1}{2^\beta - 2} + 2^\beta = 0. \tag{4.17}$$

Let $A = \frac{3^\beta - 2^\beta - 1}{2^\beta - 2}$. The discriminant $\varDelta$ and the roots of (4.17) are

$$\varDelta = A^2 - 2^\beta, \quad n_{1,2} = A \pm \sqrt{\varDelta}. \tag{4.18}$$

So $n_2 - n_1 = 2\sqrt{\varDelta}$. The interval (n_1, n_2) contains at least one positive integer n if $2\sqrt{\varDelta} > 1$ or equivalently $\varDelta > 1/4$.

We shall prove that $\varDelta > 1$ and that will finish the proof of (i). Remark that $A > (3/2)^\beta - 1$. Write

$$\varDelta = 2^\beta\left[\left(\left(\frac{3}{2\sqrt{2}}\right)^\beta - \left(\frac{1}{\sqrt{2}}\right)^\beta\right)^2 - 1\right]. \tag{4.19}$$

The function $\beta \mapsto \left(\frac{3}{2\sqrt{2}}\right)^\beta - \left(\frac{1}{\sqrt{2}}\right)^\beta$ is *increasing* because the first term increases and the second one decreases with increasing β. Hence the function

$$\beta \to \left(\left(\frac{3}{2\sqrt{2}}\right)^\beta - \left(\frac{1}{\sqrt{2}}\right)^\beta\right)^2 - 1 \tag{4.20}$$

increases, too. For $\beta = 4$, the value of (4.19) is $1/64 > 0$. As $\beta \to 2^\beta$ is also increasing, the product Δ of (4.20) times 2^β is also increasing. In short, $\beta > 5 \Rightarrow \Delta > 1$, so there exist positive integers between the two roots.

If $r = 1/2$ or $\beta = 2$, the inequality $E_1E_1E_r>$ is true because then we can write

$$D = 2(S_x - S_y)^2, \quad \text{where } S_x = \sum_{i=1}^n x_i, \quad S_y = \sum_{i=1}^n y_i.$$

If $S_x = S_y$, we have equality.

(ii) According to Theorem 2.9, the inequalities $E_pE_pE_r<$ and $E_{-p}E_{-p}E_{-r}>$ are equivalent. So if $p > r$, then $E_pE_pE_r>$ holds precisely when $E_{-p}E_{-p}E_{-r}<$ holds. If, in (i), we choose $p = 1$ and $r \geqslant 1$, then $SSE_r<$ holds. This proves the first part of (iii). If $p = -1$, we get from (i) that $E_{-1}E_{-1}E_r<$ holds for $r \geqslant 0$. So its dual (defined before Theorem 2.9) $SSE_r>$ holds if $r \leqslant 0$. (iv) Restates previous results. □

Corollary 4.4. *The inequalities PPS<, PPA<, SSA<, PPI>, SSP>, SSI> hold.*

Proof. *PPS* is $E_0E_0E_1$; *PPA* is a limiting case of $E_0E_0E_r$, $r \to \infty$ (apply Theorem 2.7(i)); *SSA* is a limiting case of $E_1E_1E_r$, $r \to \infty$; *PPI* is the dual of *PPA*; *SSP* is $E_1E_1E_0$; and *SSI* is a limiting case of $E_1E_1E_{-r}$, $r \to \infty$. □

We generalize Theorem 4.3(iii) by replacing the first summation S with a quadratic form Q. The proof follows the same lines as before.

Theorem 4.5. *If $r \geqslant 1$, then $QSE_r<$ holds. If $r \leqslant 0$, then $QSE_r>$ holds.*

Proof. Instead of (4.8), define, for any $n \geqslant 1$ and any $t \in \Re^n$,

$$D_f(x, y, t) = \sum_{1\leqslant i,j\leqslant n} \big(f(x_i + x_j) + f(y_i + y_j) - f(x_i + y_j) - f(y_i + x_j)\big)t_it_j. \tag{4.21}$$

Let $\mathbf{C}_1$ be the set of all functions $f : (0, \infty) \to \Re$ such that $D_f(x, y, t) \geqslant 0$, $\forall n \geqslant 1$, $\forall x, y \in (0, \infty)^n$, $\forall t \in \Re^n$. Then $\mathbf{C}_1$ has property (A) (Proposition 8.3) and hence contains the functions $x \to x^\beta$, $\beta = 1/r$ and $x \to -\log x$ (Corollary 8.2). The proof is the same as in Theorems 4.1 and 4.2. □

5. Inequalities of the form *SEF*

Our proof will be roughly as follows.

$$\begin{array}{ll} SPA & \text{false} \\ SPS & \text{easy} \\ SSP & \text{easy} \\ SIP & \text{from Theorem 5.3} \\ SAP & \text{from Corollary 5.4} \\ SIS & \text{from Theorem 5.5} \\ SIA & \text{from } GIA \text{ (Theorem 5.6)} \\ SSI & \text{from } GSI \text{ (Theorem 5.11)} \\ SPI & \text{from } GPI \text{ (Theorem 5.12)} \\ SAS & \text{from } SIS \text{ (Theorem 5.5)} \\ SAI & \text{from } SIA \text{ (Theorem 5.6)} \\ SSA & \text{from } GSA \text{ (Theorem 5.7)} \end{array} \tag{5.1}$$

Theorem 2.1 guarantees that *QEF* implies *SEF*. As a partial converse, we shall show that, whenever *SEF* holds with $E, F \in \{I, P, S, A\}$, then the corresponding *QEF* holds, too. We were not able to find counterexamples to this:

Conjecture. *Let $a, b \in [0, \infty)^n$. If inequality SE_qE_r holds, then the corresponding QE_qE_r holds, too.*

The 12 non-trivial *SEF* cases have $E, F \in \{I, P, S, A\}$ but $E \neq F$. From (2.58),

$$SSA \leftrightarrow SSI, \quad SIS \leftrightarrow SAS, \quad SIA \leftrightarrow SAI. \tag{5.2}$$

The other six non-trivial cases *SEF* are

$$SPA, \quad SPS, \quad SSP, \quad SPI, \quad SIP, \quad SAP. \tag{5.3}$$

The first three of these (*SPA*, *SPS*, *SSP*) are easy to handle.

We first show that $SPA< = E_1E_0E_\infty<$ is false. It is the only *SEF* case that is false. Inequality $SPA<$ states that

$$\sum_{i,j} \max(a_i, a_j) \max(b_i, b_j) \leqslant \sum_{i,j} \max(a_i, b_j) \max(a_j, b_i). \tag{5.4}$$

When $a_i = x \geqslant 0$ and $b_i = y \geqslant 0$, for all $i = 1, \ldots, n$, (5.4) states that $xy \leqslant [\max(x, y)]^2$, which is true; further, strict inequality holds if $0 \leqslant x < y$. For $n = 3$, there are many counterexamples to (5.4). For instance, if

$a = (0, 1, 2)$	$b = (1, 2, 0)$	left side = 18	right side = 17
$(1, 2, 3)$	$(2, 3, 1)$	53	52
$(2, 3, 4)$	$(3, 4, 2)$	106	105

The assertion $QPA<$ is false too. By the third example of falsehood of SPA,

$$U = \begin{pmatrix} 6 & 12 & 12 \\ 12 & 12 & 16 \\ 12 & 16 & 8 \end{pmatrix}, \quad V = \begin{pmatrix} 9 & 12 & 8 \\ 12 & 16 & 12 \\ 8 & 12 & 16 \end{pmatrix}.$$

The eigenvalues of U are -6.4317, -3.3382, 35.7699 and the eigenvalues of V are -0.0819, 5.4888, 35.5931. Therefore $\rho(U) = 35.7699 > 35.5931 = \rho(V)$.

For $n = 2$, the situation is different.

Proposition 5.1. *The inequality $QPA<$ holds for $n = 2$. Therefore $SPA<$ and $RPA<$ hold for $n = 2$. (As we indicated in Section 1, wherever P appears, we assume non-negative a, b.)*

Proof. The inequality $QPA<$ holds if $W = V - U$ is positive semidefinite, or equivalently $w_{1,1} \geqslant 0$ and $\det(W) \geqslant 0$. As $w_{1,1} = (a_1 \vee b_1)^2 - a_1 b_1 = (a_1 \vee b_1)|a_1 - b_1| \geqslant 0$, one has only to check the second condition,

$$\begin{aligned} &[(a_1 \vee b_1)|a_1 - b_1|][(a_2 \vee b_2)|a_2 - b_2|] \\ &- [(a_1 \vee b_2)(a_2 \vee b_1) - (a_1 \vee a_2)(b_1 \vee b_2)]^2 \geqslant 0. \end{aligned} \tag{5.5}$$

Let

$$\begin{aligned} A &= [(a_1 \vee b_1)|a_1 - b_1|][(a_2 \vee b_2)|a_2 - b_2|], \\ B &= [(a_1 \vee b_2)(a_2 \vee b_1) - (a_1 \vee a_2)(b_1 \vee b_2)]^2. \end{aligned} \tag{5.6}$$

Then $\det(W) = A - B$.

Let us order the four numbers a_1, a_2, b_1, b_2 as $\alpha_1 \leqslant \alpha_2 \leqslant \alpha_3 \leqslant \alpha_4$. That can be done in $4! = 24$ ways. These 24 cases belong to three different classes.

Class 1. $A = \alpha_2\alpha_4(\alpha_2 - \alpha_1)(\alpha_4 - \alpha_3)$, $B = 0$. This class contains eight cases:

$$\begin{aligned} &a_1 \leqslant b_1 \leqslant a_2 \leqslant b_2, \quad a_1 \leqslant b_1 \leqslant b_2 \leqslant a_2, \quad b_1 \leqslant a_1 \leqslant a_2 \leqslant b_2, \\ &b_1 \leqslant a_1 \leqslant a_2 \leqslant b_2, \quad a_2 \leqslant b_2 \leqslant a_1 \leqslant b_1, \quad a_2 \leqslant b_2 \leqslant b_1 \leqslant a_1, \\ &b_2 \leqslant a_2 \leqslant a_1 \leqslant b_1, \quad b_2 \leqslant a_2 \leqslant b_1 \leqslant a_1. \end{aligned}$$

These cases may be expressed more concisely as $a_1 \vee b_1 \leqslant a_2 \wedge b_2$ or $a_2 \vee b_2 \leqslant a_1 \wedge b_1$. Obviously $A - B = A \geqslant 0$.

Class 2. $A = \alpha_3\alpha_4(\alpha_3 - \alpha_1)(\alpha_4 - \alpha_2)$, $B = {\alpha_4}^2(\alpha_3 - \alpha_2)^2$. This class contains eight cases:

$$\begin{aligned} &a_1 \leqslant a_2 \leqslant b_1 \leqslant b_2, \quad a_1 \leqslant b_2 \leqslant b_1 \leqslant a_2, \quad b_1 \leqslant a_2 \leqslant a_1 \leqslant b_2, \\ &b_1 \leqslant b_2 \leqslant a_1 \leqslant a_2, \quad a_2 \leqslant a_1 \leqslant b_2 \leqslant b_1, \quad a_2 \leqslant b_1 \leqslant b_2 \leqslant a_1, \\ &b_2 \leqslant b_1 \leqslant a_2 \leqslant a_1, \quad b_2 \leqslant a_1 \leqslant a_2 \leqslant b_1. \end{aligned}$$

Replacing α_1 with α, α_2 with $\alpha + x$, α_3 with $\alpha + x + y$, and α_4 with $\alpha + x + y + z$, where $\alpha, x, y, z \geqslant 0$, one gets

$$A - B = \alpha_4[\alpha(xy + yz + zx) + x(x + y + z) + xyz] \tag{5.7}$$

which is obviously non-negative.

Class 3. $A = \alpha_3\alpha_4(a_4 - \alpha_1)(\alpha_3 - a_2)$, $B = \alpha_4^2(\alpha_3 - \alpha_2)^2$. This class contains eight cases:

$$\begin{aligned}
&a_1 \leqslant a_2 \leqslant b_2 \leqslant b_1, \quad a_1 \leqslant b_2 \leqslant a_2 \leqslant b_1, \quad b_1 \leqslant a_2 \leqslant b_2 \leqslant a_1,\\
&b_1 \leqslant b_2 \leqslant a_2 \leqslant a_1, \quad a_2 \leqslant a_1 \leqslant b_1 \leqslant b_2, \quad a_2 \leqslant b_1 \leqslant a_1 \leqslant b_2,\\
&b_2 \leqslant b_1 \leqslant a_1 \leqslant a_2, \quad b_2 \leqslant b_1 \leqslant a_1 \leqslant a_2.
\end{aligned}$$

More concisely, the interval bounded by a_1, b_1 contains or is contained in the interval bounded by a_2, b_2. Replacing α_1 with α, α_2 with $\alpha + x$, α_3 with $\alpha + x + y$, and α_4 with $\alpha + x + y + z$, where $\alpha, x, y, z \geqslant 0$, one gets

$$A - B = \alpha_4(\alpha_3 - \alpha_2)[\alpha(x + z) + x(x + y + z)], \tag{5.8}$$

again obviously non-negative. □

Proposition 5.2. *The inequalities QPS< and QSP> hold. As a consequence, SPS< (=$E_1E_0E_1$<) and SSP> (=$E_1E_1E_0$>) and RPS< and RSP> also hold.*

Proof. *QPS<* states that

$$\sum_{i,j}(a_i + a_j)(b_i + b_j)x_ix_j \leqslant \sum_{i,j}(a_i + b_j)(b_i + a_j)x_ix_j, \quad \forall x \in \Re^n.$$

The a_ib_i terms and a_jb_j terms on each side cancel each other, leaving only

$$\sum_{i,j}(a_ib_j + a_jb_i)x_ix_j \leqslant \sum_{i,j}(a_ia_j + b_ib_j)x_ix_j,$$

which amounts to the obvious inequality $2pq \leqslant p^2 + q^2$, with $p = \sum_i a_ix_i$, $q = \sum_i b_ix_i$. Similarly, *QSP>* means that

$$\sum_{i,j}(a_ia_j + b_ib_j)x_ix_j \geqslant \sum_{i,j}(a_ib_j + b_ia_j)x_ix_j,$$

which reduces to $p^2 + q^2 \geqslant 2pq$. □

Theorem 5.3. *Inequality QIP< holds, i.e.,*

$$\sum_{i,j}\min(a_ia_j, b_ib_j)x_ix_j \leqslant \sum_{i,j}\min(a_ib_j, b_ia_j)x_ix_j, \quad \forall a, b \in [0, \infty)^n, \quad x \in \Re^n. \tag{5.9}$$

As a consequence, SIP< and RIP< also hold.

Proof. Consider the function

$$r(s,t) = \min(s,t) - \min(1, st), \quad \text{where } s,t \in [0,\infty). \tag{5.10}$$

It suffices to exhibit random variables $Z(t)$ $(t \geqslant 0)$ that satisfy

$$\mathrm{Cov}(Z(s), Z(t)) = r(s,t) \quad \text{for all } s,t \geqslant 0, \tag{5.11}$$

for then

$$\sum_{i,j} r(t_i, t_j) x_i x_j = \sum_{i,j} x_i x_j \,\mathrm{Cov}(Z(t_i), Z(t_j)) = \mathrm{var}\left(\sum_{i=1}^{n} x_i Z(t_i)\right) \geqslant 0, \tag{5.12}$$

for all choices of n, $t_i \in [0,\infty)$ and $x_i \in \Re$ $(i = 1,\ldots,n)$. Hence, using (5.10),

$$\sum_{i,j} \min(1, t_i t_j) a_i a_j \leqslant \sum_{i,j} \min(t_i, t_j) a_i a_j.$$

If $a_i > 0$, $b_i > 0$, $i = 1,\ldots,n$, then the last inequality with $t_i = b_i/a_i$ immediately yields (5.9). The case $a_i \geqslant 0$, $b_i \geqslant 0$ follows by continuity.

Now we construct random variables $Z(t)$ that satisfy (5.11). The standard Brownian motion $W(t)(t \geqslant 0;\, W(0) = 0)$ satisfies

$$\mathrm{Cov}(W(s), W(t)) = s \wedge t, \quad \forall s,t \geqslant 0.$$

Define

$$Z(t) = \begin{cases} W(t) - tW(1) & \text{if } 0 \leqslant t \leqslant 1, \\ W(1) - tW\left(\dfrac{1}{t}\right) & \text{if } t \geqslant 1. \end{cases}$$

Then $Z(0) = Z(1) = 0$ and $\{Z(t); 0 \leqslant t \leqslant 1\}$ is the usual Brownian bridge which satisfies, for $0 \leqslant s,\ t \leqslant 1$, $\mathrm{Cov}(Z(s), Z(t)) = \mathrm{Cov}(W(s) - sW(1), W(t) - tW(1)) = s \wedge t - st - st + st = s \wedge t - st = r(s,t)$. Here we used (5.10). This verifies (5.11) when $0 \leqslant s,t \leqslant 1$. When $s,t \geqslant 1$, then

$$\begin{aligned}
\mathrm{Cov}(Z(s), Z(t)) &= \mathrm{Cov}\left(-sZ\left(\frac{1}{s}\right), -tZ\left(\frac{1}{t}\right)\right) \\
&= st\mathrm{Cov}\left(Z\left(\frac{1}{s}\right), Z\left(\frac{1}{t}\right)\right) \\
&= st\left[\left(\frac{1}{s} \wedge \frac{1}{t}\right) - \frac{1}{st}\right] = s \wedge t - 1 = r(s,t).
\end{aligned}$$

Finally, if $0 \leqslant s \leqslant 1 \leqslant t$, then

$$\begin{aligned}
\mathrm{Cov}(Z(s), Z(t)) &= \mathrm{Cov}\left(Z(s), -tZ\left(\frac{1}{t}\right)\right) = -t\mathrm{Cov}\left(Z(s), Z\left(\frac{1}{t}\right)\right) \\
&= -t\left[\min\left(s, \frac{1}{t}\right) - \frac{s}{t}\right] = s - \min(st, 1) = r(s,t). \qquad \square
\end{aligned}$$

Remark. To see whether (5.9) and (5.12) hold with equality, apply the following operations (where n may be replaced by a smaller integer).

(i) Ignore all indices i for which either $a_i = 0$ or $b_i = 0$.
(ii) Lump together into a single index ρ all indices i with the same ratio $b_i/a_i = t$ letting $a_\rho = \sum_{\frac{b_i}{a_i}=t} a_i$ and $b_\rho = \sum_{\frac{b_i}{a_i}=t} b_i$, for all $t > 0$. The new pair a_ρ, b_ρ has the same values as the old pair for the left side and right side of (5.9).
(iii) Permute indices to satisfy (5.16) in the proof below.

Then the old pair a, b satisfies (5.9) with the equality sign if and only if the new vector b is the reversal of the new vector a, where reversal is defined in the proof below after (5.17).

Proof. Set $Y = \sum_{i=1}^n a_i Z(t_i)$. Then $EY = 0$. Equality holds in (5.12) if and only if $Y = 0$ with probability 1. Let $t_i = b_i/a_i$ and let $\{\tau_1, \ldots, \tau_k\}$ be the distinct numbers among $\{t_i; a_i \neq 0; 0 < t_i < 1, i = 1, \ldots, n\} \cup \{1/t_i; a_i \neq 0; t_i > 1, i = 1, \ldots, n\}$. One may assume that $0 < \tau_1 < \cdots < \tau_k < 1$. Since $Z(0) = Z(1) = 0$, Y is a linear combination of the form $Y = \sum_{r=1}^k c_r Z(\tau_r)$. Because Brownian motion $W(t)$ is a process of independent random increments, $Y = 0$ if and only if $c_r = 0, r = 1, \ldots, k$. Equivalently, $Y = 0$ if and only if

$$\sum_{t_i=\tau_r} a_i = \frac{1}{\tau_r} a_i \quad \text{for } r = 1, \ldots, k. \tag{5.13}$$

This allows us to determine when (5.9) holds with the equality sign. If, for instance, $a_r = 0$, then the terms with $i = r$ or $j = r$ are always zero and can be ignored. If $a_r = b_r$, then each term $u_{i,j}$ on the left side of (5.9) with $i = r$ or $j = r$ exactly cancels the corresponding term $v_{i,j}$ on the right side of (5.9). Replacing n by a smaller integer if necessary, we may assume that

$$a_i > 0, \quad b_i > 0, \quad a_i \neq b_i \quad \text{for all } i = 1, \ldots, n. \tag{5.14}$$

Then $t_i = b_i/a_i$ satisfies $t_i \neq 0$, $t_i \neq 1$, $i = 1, \ldots, n$. Condition (5.13) for equality becomes

$$\sum_{\frac{b_i}{a_i}=\tau_r} a_i = \sum_{\frac{a_i}{b_i}=\tau_r} b_i \quad \text{for } r = 1, \ldots, k. \tag{5.15}$$

After lumping as prescribed in part (ii) above, we may assume that all the ratios $t_i = b_i/a_i$ are different. Permuting indices, we may assume that

$$0 < \frac{b_1}{a_1} < \frac{b_2}{a_2} < \cdots < \frac{b_r}{a_r} < 1 < \frac{b_{r+1}}{a_{r+1}} < \cdots < \frac{b_n}{a_n}. \tag{5.16}$$

Then the necessary and sufficient condition (5.15) for equality can hold only when $n = 2m$ is even, and when $r = m = n/2$, and finally when

$$a_j = b_{2m+1-j}, \quad b_j = a_{2m+1-j}, \qquad j = 1, \ldots, m. \tag{5.17}$$

This is the same as $b_j = a_{2m-j}$ for all $j = 1, \dots, n = 2m$. Thus (assuming (5.16)), the vector b is a *reversal* of the vector a, that is, $b_{\sigma(i)} = a_i$ $(i = 1, \dots, n)$ where the permutation σ is its own inverse. The first Remark in Section 2 guarantees that in such a case $DU = DV$. $\square$

Example. Inequality (5.9) holds with the equality sign if $n = 8$ and

$$a = (p, q, 3r, 3s, 2t, 2u, 3v, 3w), \quad b = (3p, 3q, r, s, 3t, 3u, 2v, 2w). \tag{5.18}$$

Here $p, q, \dots, v, w$ are positive and such that $p + q = r + s$; $t + u = v + w$.

Remark. Here is a proof by induction of *SIP*<. The function r defined at (5.10) satisfies $r(s, t) = -t \times r(s, 1/t)$, $r(s, 0) = r(s, 1) = 0$, $r(t, t) = t(1 - t)$ if $0 \leqslant t \leqslant 1$ while $r(t, t) = t - 1$ if $t \geqslant 1$. Let $t_i = b_i/a_i$ and, as at (5.12), let

$$Q_n(t, x) = \sum_{i,j} r(t_i, t_j)x_i x_j, \; x \in \Re^n. \tag{5.19}$$

We will prove that always $Q_n(t, x) \geqslant 0$. The proof is by induction with respect to n. The case $n = 1$ is trivial since $r(t, t) \geqslant 0$. Let $n \geqslant 2$ be fixed.

Lemma. *For any $t \in [0, \infty)^n$ and $x \in \Re^n$ and $i, j \in \{1, \dots, n\}$, $Q_n(t, x) \geqslant 0$ whenever one of the following occurs*:

(i) *Either $t_i = 0$ for some i or $t_j = 1$ for some j.*
(ii) *$t_i = t_j$ for some i, j with $i \neq j$.*
(iii) *$t_i t_j = 1$ for some i, j with $i \neq j$.*

Such points $t \in [0, \infty)^n$ will be said to be "special points".

Proof. (i) Let (for example) $t_1 = 0$ or $t_1 = 1$. Since $r(t_1, t_j) = 0$ for all *j*,

$$Q_n(t, x) = \sum_{i=2}^{n} \sum_{j=2}^{n} r(t_i, t_j)x_i x_j \geqslant 0, \tag{5.20}$$

where the inequality holds by induction.

(ii) Suppose $t_1 = t_2$ (say). Let $\tau_1 = t_1, \tau_2, \dots, \tau_m$ be the distinct values of $t_1, \dots, t_n$.

Let $c_p \geqslant 1$ denote the number of t_j equal to τ_p $(p = 1, \dots, m)$. Then $c_1 \geqslant 2$ and $c_1 + \cdots + c_m = n$, thus $m < n$. Let

$$J_p = \{j \in \{1, \dots, n\}; t_j = \tau_p\}. \tag{5.21}$$

Thus $|J_p| = c_p$; and put $\alpha_p = \sum\{x_j;\ j \in J_P\}$, $p = 1, \dots, m$. Then, as is easily seen,

$$Q_n(t,x) = \sum_{p=1}^{m}\sum_{q=1}^{m} r(\tau_p,\tau_q)\alpha_p\alpha_q = Q_m(\tau,\alpha) \geqslant 0,$$

by induction.

(iii) Suppose (for concreteness) that $0 < t_1 < 1 < t_2$ satisfy $t_1t_2 = 1$. Let $\alpha = (\alpha_2,\ldots,\alpha_n)$ be defined by $\alpha_2 = a_2 - t_1a_1$ while $\alpha_j = a_j$ for all $3 \leqslant j \leqslant n$. Putting $t^1 = (t_2,\ldots,t_n)$, we have, by induction, that

$$Q_{n-1}(t^1,\alpha) = \sum_{i=2}^{n}\sum_{j=2}^{n} r(t_i,t_j)\alpha_i\alpha_j \geqslant 0.$$

It suffices to show that $Q_n(t,x) = Q_{n-1}(t^1,\alpha)$. The terms $r(t_i,t_j)a_ia_j$ with $3 \leqslant i, j \leqslant n$ in $Q_n(t,a)$ and $Q_{n-1}(t^1,\alpha)$ cancel each other. Thus, it suffices to show that $B_1 + 2B_2 = C_1 + 2C_2$ where

$$\begin{aligned}
B_1 &= r(t_1,t_1)a_1^2 + r(t_2,t_2)a_2^2 + 2r(t_1,t_2)a_1a_2,\\
B_2 &= \sum_{j=3}^{n}(r(t_1,t_j)a_1a_j + r(t_2,t_j)a_2a_j),\\
C_1 &= r(t_2,t_2)\alpha_2^2 = r(t_2,t_2)(a_2 - t_1a_1)^2,\\
C_2 &= \sum_{j=3}^{n} r(t_2,t_j)\alpha_2a_j = \sum_{j=3}^{n} r(t_2,t_j)(a_2 - t_1a_1)a_j.
\end{aligned}$$

Here, $r(t_1,t_1) = t_1^2 r(t_2,t_2)$, $r(t_1,t_j) = -t_1r(t_2,t_j)$ for all $j \geqslant 2$, implying $B_1 = C_1$ and $B_2 = C_2$. This settles case (iii) and thus the lemma. $\square$

Let $x \in \Re^n$ be fixed and let $t \in [0,\infty)^n$. From our lemma, $Q_n(t^0,a) < 0$ would imply that $t = t^0$ is not a special point. Thus, we may assume that

$$t_j^0 \neq 0,\quad t_j^0 \neq 1,\quad t_i^0t_j^0 \neq 1 \quad \text{for all } i, j, \quad t_i^0 \neq t_j^0 \ \text{ if } i \neq j.$$

Permuting indices, $0 < t_1^0 < t_2^0 < \cdots < t_n^0$. Let us now study $Q_n(t,a)$ as a function of the first coordinate $y = t_1$, with y close to $y_0 = t_1^0$. The other coordinates are fixed, thus $t_2 = t_2^0,\ldots,t_n = t_n^0$. One has $Q_n(t,a) = \phi(y)$, where

$$\phi(y) := \psi(y) + 2\sum_{j=2}^{n}\left(\min\left(y,t_j^0\right) - \min\left(1,yt_j^0\right)\right)a_1a_j + C. \tag{5.22}$$

Here C is independent of $t_1 = y$ while $\psi(y) := r(y,y)a_1^2 = y(1-y)a_1^2$ if $0 \leqslant y \leqslant 1$; $\psi(y) = (y-1)a_1^2$ if $y \geqslant 1$. $\psi(y)$ is concave on [0, 1] and linear on $[1,\infty]$. The sum in (5.22) represents a linear function of y as long as the interval (y, y_0) does not contain any of the values 1, t_j^0 and $1/t_j^0$ (which values are all different from $y^0 = t_1^0$ since $y^0 = t^0$ is non-special). This range of y amounts to a closed interval $[y_1, y_2]$ with $0 \leqslant y_1 < y_0 < y_2$. Further, $y_2 \leqslant t_2^0 < +\infty$.

If $t^0 > 1$, then $\psi(y)$ is linear in the interval $[y_1, y_2]$. If $t_1^0 < 1$, then $\psi(y)$ is concave on $[y_1, y_2]$. In each case, $\psi(y)$ and thus $\phi(y)$ are concave on $[y_1, y_2]$.

The extreme points t^1 and $t = t^2$ in $[0, \infty)^n$ associated with $y = y_1$ and $y = y_2$ are clearly special points. For there $t_1 = y$ must necessarily take one of the values $0, 1, t_j^0$ or $1/t_j^0$, where $j \geqslant 2$. From the above lemma, both $\phi(y_1) \geqslant 0$ and $\phi(y_2) \geqslant 0$. Hence, the concavity of $\phi(y)$ on $[y_1, y_2]$ yields that $\phi(y) \geqslant 0$ for all $y_1 \leqslant y \leqslant y_2$. Hence $\phi(y_0) = \phi(t_1^0) = Q_n(t^0, a) < 0$ is impossible. $\square$

Corollary 5.4. *The inequalities QAP>, that is,*

$$\sum_{i,j} \max(a_i a_j, b_i b_j) x_i x_j \geqslant \sum_{i,j} \max(a_i b_j, b_i a_j) x_i x_j,$$
$$\forall n \geqslant 1, \quad \forall a, b \in [0, \infty)^n, \quad x \in \Re^n, \tag{5.23}$$

SAP> and RAP> all hold.

Proof. We will use *QIP<*, that is, inequality (5.9). abbreviated to $L_1 \leqslant R_1$, to prove (5.23), abbreviated to $L_2 \geqslant R_2$. It suffices to show that $L_2 - R_2 \geqslant R_1 - L_1$, that is, $L_1 + L_2 \geqslant R_1 + R_2$. Since $x \wedge y + x \vee y = x + y$, the latter is equivalent to

$$\sum_{i,j} (a_i a_j + b_i b_j) x_i x_j \geqslant \sum_{i,j} (a_i b_j + b_i a_j) x_i x_j.$$

This is precisely the trivial inequality *SSP*, saying that $p^2 + q^2 \geqslant 2pq$, where $p = \sum a_i x_i$ and $q = \sum b_i x_i$. $\square$

Theorem 5.5. *The inequalities QIS<, that is,*

$$\sum_{i,j} ((a_i + a_j) \wedge (b_i + b_j)) x_i x_j \leqslant \sum_{i,j} ((a_i + b_j) \wedge (a_j + b_i)) x_i x_j,$$
$$\forall n \geqslant 1, \quad a, b \in [0, \infty)^n, \quad x \in \Re^n, \tag{5.24}$$

SIS<, SAS>, PAP>, PIP< and RIS< (this last one if $a, b > 0$) all hold.

Proof. From (2.60), *PIP* $\leftrightarrow$ *PAP* $\leftrightarrow$ *SIS* $\leftrightarrow$ *SAS*, so it suffices to prove (5.24). In *QIP<* (5.9), replace a_i by $a_i + \lambda$ and b_i by $b_i + \lambda$ $(1 \leqslant i \leqslant n)$, where $\lambda \geqslant 0$. The n^2 terms λ^2 on each side cancel each other. Hence

$$\sum_{i,j} (a_i a_j + \lambda(a_i + a_j)) \wedge (b_i b_j + \lambda(b_i + b_j)) x_i x_j$$
$$\leqslant \sum_{i,j} (a_i b_j + \lambda(a_i + b_j)) \wedge (a_j b_i + \lambda(a_j + b_i)) x_i x_j.$$

Divide both sides by λ, and let $\lambda \to \infty$. $\square$

Theorem 5.6. *For every non-decreasing $f : J \to \Re$ where J is an interval containing all the a_i and b_i, the generalized inequality*

$$\sum_{i,j} f((a_i \vee a_j) \wedge (b_i \vee b_j))x_i x_j \leqslant \sum_{i,j} f((a_i \vee b_j) \wedge (b_i \vee a_j))x_i x_j, \quad \forall n \geqslant 1, \;\; \forall x \in \Re^n \qquad (GIA<)$$

holds. Hence the inequalities QIA<, namely,

$$\sum_{i,j}((a_i \vee a_j) \wedge (b_i \vee b_j))x_i x_j \leqslant \sum_{i,j}((a_i \vee b_j) \wedge (b_i \vee a_j))x_i x_j, \quad \forall n \geqslant 1, \;\; \forall a, b, x \in \Re^n, \qquad (5.25)$$

and SIA<, SAI>, PIA<, PAI> and RIA< are true, the last one if $a, b \geqslant 0$. Equality holds in (GIA<) for every non-decreasing function f if and only if, for every $c \in \Re$, the number of those i such that $a_i \geqslant c$ is the same as the number of those i such that $b_i \geqslant c$, that is, if there exists a permutation σ such that $b_i = a_{\sigma(i)}$.

Proof. Recall from (2.60) that $PIA \leftrightarrow PAI \leftrightarrow SIA \leftrightarrow SAI$.

Step 1. Prove $GIA<$ for $f = 1_{[c,\infty)}$, where $c \in \Re$.

Let $N = \{1, 2, \ldots, n\}$ and think of x as a signed measure on N with the weights x_i. So $x(A)$ means $\sum_{i \in A} x_i$ and if $C \subset N \times N$ then $x \otimes x(C) = \sum_{(i,j)\in C} x_i x_j$.

Let $D = \sum_{i,j}[f((a_i \vee b_j) \wedge (a_j \vee b_i)) - f((a_i \vee a_j) \wedge (b_j \vee b_i))]x_i x_j$ be the difference between the right and left sides of $(GIA<)$. We shall prove that $D \geqslant 0$. Let

$$\omega_{i,j} = f((a_i \vee b_j) \wedge (a_j \vee b_i)) - f((a_i \vee a_j) \wedge (b_j \vee b_i)). \qquad (5.26)$$

Then $\omega_{i,j} \in \{-1, 0, 1\}$. Precisely, $\omega_{i,j} = 1 \Leftrightarrow (a_i \vee b_j) \wedge (b_i \vee a_j) \geqslant c$ and $(a_i \vee a_j) \wedge (b_i \vee b_j) < c$. Let $A = \{i \in N; a_i \geqslant c\}$ and $B = \{i \in N; b_j \geqslant c\}$. As $(a_i \vee b_j) \wedge (b_i \vee a_j) \geqslant c \Leftrightarrow a_i \vee b_j \geqslant c$ and $b_i \vee a_j \geqslant c \Leftrightarrow (i \in A$ or $j \in B)$ and $(i \in B$ or $j \in A) \Leftrightarrow ((i, j) \in A \times N$ or $(i, j) \in N \times B)$ and $((i, j) \in B \times N$ or $(i, j) \in N \times A)$, we see that

$$(a_i \vee b_j) \wedge (b_i \vee a_j) \geqslant c \iff (i, j) \in AB \times N \cup A \times A \cup B \times B \cup N \times AB. \qquad (5.27)$$

Similarly, $(a_i \vee a_j) \wedge (b_i \vee b_j) < c \Leftrightarrow a_i \vee a_j < c$ or $b_i \vee b_j < c \Leftrightarrow (a_i < c$ and $a_j < c)$ or $(b_i < c$ and $b_j < c) \Leftrightarrow (i, j) \in (A^c \times N \cap N \times A^c) \cup (B^c \times N \cap N \times B^c)$. Therefore

$$(a_i \vee a_j) \wedge (b_i \vee b_j) < c \iff (i, j) \in (A^c \times A^c) \cup (B^c \times B^c). \qquad (5.28)$$

Combining (5.27) and (5.28) gives

$$\omega_{i,j} = 1 \iff (i, j) \in (AB^c \times AB^c) \cup (BA^c \times BA^c) := C_1. \qquad (5.29)$$

On the other hand, $\omega_{i,j} = -1 \Leftrightarrow (a_i \vee b_j) \wedge (b_i \vee a_j) < c$ and $(a_i \vee a_j) \wedge (b_i \vee b_j) \geqslant c$. Similar considerations yield

$$\omega_{i,j} = -1 \iff (i, j) \in (AB^c \times BA^c) \cup (BA^c \times AB^c) := C_2. \tag{5.30}$$

As a consequence, $D = \sum_{i,j} \omega_{i,j} x_i x_j = \sum_{(i,j)\in C_1} x_i x_j - \sum_{(i,j)\in C_2} x_i x_j = x \otimes x(C_1) - x \otimes x(C_2) = x(AB^c)^2 + x(BA^c)^2 - 2x(AB^c)x(BA^c) = (x(AB^c) - x(BA^c))^2 \geqslant 0$ and we are done.

Step 2. Let us denote by $\mathbf{H}$ the set of those functions $f : J \to \Re$ for which *GIA<* holds. $\mathbf{H}$ is a positive cone, that is, $f, g \in \mathbf{H}$, $\lambda, \mu \geqslant 0 \Rightarrow \lambda f + \mu g \in \mathbf{H}$. Moreover, $\mathbf{H}$ is sequentially closed, that is, if $(f_m)_m$ is a sequence of functions from $\mathbf{H}$ such that $f_m \to f$, then $f \in \mathbf{H}$. By Step 1, $\mathbf{H}$ contains the functions of the form $f = 1_{[c,\infty)}$. Any non-decreasing function f is the limit of a positive linear combination of functions of this type, hence any non-decreasing function f belongs to $\mathbf{H}$. Thus *GIA<* holds for any non-decreasing function. □

Remark. All that matters are the values of f on the finite set $\{a_i \vee a_j, a_i \vee b_j, b_i \vee b_j; 1 \leqslant i, j \leqslant n\}$, so any function f behaves as a step function.

We prove next a generalization of *SSA<* which will be used in the next section. It also gives an alternative proof of *SSI>*.

Theorem 5.7. *Let $J = (\theta_0, \theta_1)$ be an interval and $f : J + J = (2\theta_0, 2\theta_1) \to \Re$ be concave and non-decreasing on J. Let $n \geqslant 1$, $a, b \in J^n$ and $x \in \Re^n$. Then*

$$\sum_{1\leqslant i,j\leqslant n} f(a_i \vee a_j + b_i \vee b_j) x_i x_j \leqslant \sum_{1\leqslant i,j\leqslant n} f(a_i \vee b_j + b_i \vee a_j) x_i x_j. \tag{$GSA<$}$$

Proof. Let $v_{i,j} = a_i \vee b_j + a_j \vee b_i$ and $u_{i,j} = a_i \vee a_j + b_j \vee b_i$. *GSA<* holds if and only if $S = S(a, b, x) \geqslant 0$, $\forall a, b, x$ where

$$S(a, b, x) = \sum_{1\leqslant i,j\leqslant n} (f(v_{i,j}) - f(u_{i,j})) x_i x_j. \tag{5.31}$$

The proof of (5.31) uses induction on the number k of different values of a_i and b_j. Denote by $Z = Z(a, b) = \{a_i, b_i; 1 \leqslant i \leqslant n\}$. We order these different values as $Z = \{\theta_0 < z_1 < \cdots < z_k < \theta_1\}$.

If $k = 1$, (5.31) is obvious because then $u_{i,j} = v_{i,j}$ and $S = 0$.

Let n, x be fixed and $k = |Z| \geqslant 2$. The induction assumption is that $S(a', b') \geqslant 0$ for any $a', b' \in J^n$ such that $|Z(a, b)| < k$.

Let $t \in (\theta_0, z_2)$. Let $a(t)$ denote the vector contained in J^n obtained from a by replacing a_i by t each time $a_i = z_1$. We construct $b(t)$ similarly: each time $b_i = z_1$ we replace b_i with t.

Then $Z(a(t), b(t)) = \{t < z_2 < \cdots < z_k\}$ and $Z(a(z_2), b(z_2)) = \{z_2 < \cdots < z_k\}$. Thus $|Z(a(t), b(t))| = k$ for $t \neq z_2$ and $|Z(a(z_2), b(z_2))| = k - 1$.

Let $S(t)$ denote the sum from (5.31) with a, b replaced with $a(t), b(t)$. Then $S(z_2) \geqslant 0$ according to our induction assumption. If we prove that the function $S(t)$ is non-increasing on $\theta_0 < t < z_2$, then we are done. We shall derive an explicit formula for $S(t)$.

Recall that $x \in \Re^n$ will not be changed. As in the proof of *GIA<*, we interpret x as a signed measure on $N = \{1, 2, \ldots, n\}$. Let $1 \leqslant r, s \leqslant k$ and

$$A_{r,s} = \{i;\, 1 \leqslant i \leqslant n \text{ and } a_i = z_r, b_i = z_s\}, \quad m_{r,s} = x(A_{r,s}) = \sum_{i \in A_{r,s}} x_i. \tag{5.32}$$

There are k^2 sets $A_{r,s}$, some of them possibly empty. They are disjoint and their union is $N = \{1, 2, \ldots, n\}$. For the values of $A_{r,s}$ and $m_{r,s}$ associated with the pair $(a(t), b(t))$, when z_1 is replaced by t, we write $A_{r,s}(t), m_{r,s}(t)$.

Consider $i, j \in N$ such that $i \in A_{r,s}$ and $j \in A_{p,q}$. This means that $a_i = z_r, b_i = z_s, a_j = z_p$ and $b_j = z_q$. Moreover, $u_{i,j} = a_i \vee a_j + b_j \vee b_i = z_r \vee z_p + z_s \vee z_q = z_{p\vee r} + z_{q\vee s}$ and similarly $v_{i,j} = z_{q\vee r} + z_{p\vee s}$. Replacing these quantities in (5.31) yields

$$S(t) = \sum_{r,s} \sum_{p,q} [f(z_{q\vee r} + z_{p\vee s}) - f(z_{p\vee r} + z_{q\vee s})] m_{r,s}(t) m_{p,q}(t). \tag{5.33}$$

Here and throughout, each z_1 is to be replaced with t. Let $g(t)$ denote the sum of all terms in (5.33) that involve t. To prove that $S(t)$ is non-increasing, it suffices to prove it for $g(t)$. We claim that if $\theta_0 < t < z_2$ (as assumed above), then

$$g(t) = c_2 f(t + z_2) + c_3 f(t + z_3) + \cdots + c_k f(t + z_k) \tag{5.34}$$

where

$$\begin{aligned} c_m = {} & \sum_{p\vee s=m} m_{1,s} m_{p,1} + \sum_{q\vee r=m} m_{1,q} m_{r,1} \\ & - \left(\sum_{q\vee s=m} m_{1,q} m_{1,s} + \sum_{p\vee r=m} m_{p,1} m_{r,1} \right). \end{aligned} \tag{5.35}$$

For instance, the first sum on the right side of (5.35) derives from the fact that $f(z_{q\vee r} + z_{p\vee s}) = f(t + z_m)$ if $q = r = 1$ and $p \vee s = m$; similarly for the other three sums on the right side of (5.35). Keeping the first sum in (5.35) as it is and renaming the summation variables in the other three sums, (5.35) simplifies to

$$c_m = -\sum_{p\vee s=m} \varDelta_s \varDelta_p \quad \text{where } \varDelta_s = m_{s,1} - m_{1,s}. \tag{5.36}$$

In particular, $\varDelta_1 = 0$, thus $c_1 = 0$, agreeing with the fact that, in (5.33), the term $f(t + t)$ has coefficient 0. Then (5.36) implies that

$$c_1 + c_2 + \cdots + c_m = - \sum_{p,s:1 \leqslant p \vee s \leqslant m} \Delta_s \Delta_p = -(\Delta_1 + \cdots + \Delta_m)^2 \leqslant 0.$$

Let $\sigma_m = (\Delta_1 + \cdots + \Delta_m)^2$. Remarking that $c_m = \sigma_{m-1} - \sigma_m$ we deduce from (5.34) and the fact that $\sigma_1 = 0$ that

$$\begin{aligned} g(t) &= \sum_{m=2}^{k} (\sigma_{m-1} - \sigma_m) f(t + z_k) \\ &= \sum_{m=2}^{k-1} \sigma_m [f(t + z_{m+1}) - f(t + z_m)] - \sigma_k f(t + z_k). \end{aligned}$$

As $\sigma_m \geqslant 0$, $\forall m$, this latter formula clearly shows that g is decreasing: the last term is decreasing since f is increasing and the rest of terms are all decreasing since any concave function f has the property that $x \mapsto f(x + b) - f(x + a)$ is decreasing if $b > a$. In our case $a = z_k$ and $b = z_{k+1}$. □

Corollary 5.8. *Inequality QSA< holds; therefore SSA< holds. If $a, b \geqslant 0$, then RSA< holds.*

Proof. Take $f(x) = x$. It is concave and increasing. □

Corollary 5.9. *Inequality PSA< holds.*

Proof. Take $f(x) = \log x$. It is concave and increasing. □

Theorem 5.10. *Inequalities QSI> and QSA<, explicitly*

$$\sum_{i,j} ((a_i \wedge a_j) + (b_i \wedge b_j)) x_i x_j \geqslant \sum_{i,j} ((a_i \wedge b_j) + (b_i \wedge a_j)) x_i x_j \tag{5.37}$$

and

$$\sum_{i,j} ((a_i \vee a_j) + (b_i \vee b_j)) x_i x_j \leqslant \sum_{i,j} ((a_i \vee b_j) + (b_i \vee a_j)) x_i x_j, \tag{5.38}$$

hold, and so do PPI>, PPA<, SSA<.

Proof. The proof of *QSI>* is similar to the proof of *QIP<*. One has to check that the matrix W is semipositive definite, where $w_{i,j} = [a_i \wedge a_j + b_i \wedge b_j] - [a_i \wedge b_j + a_j \wedge b_i]$. Let $W(t)$ denote the usual Brownian motion with covariance $\mathrm{Cov}(W(s), W(t)) = s \wedge t$ and let $Z_i = W(a_i) - W(b_i)$. The vector $Z = (Z_i)_{1 \leqslant i \leqslant n}$ has the covariance W. That proves (5.37). The W matrix is the same for *QSI>* and *QSA<*, since $x \wedge y + x \vee y = x + y$. This proves (5.38). In view of *PPI* $\leftrightarrow$ *PPA* $\leftrightarrow$ *SSI* $\leftrightarrow$ *SSA* from (2.60), *PPI>*, *PPA<*, *SSA<* hold. □

Theorem 5.11. *The generalized SI inequality*

$$\sum_{i,j} g((a_i \wedge a_j) + (b_i \wedge b_j))x_i x_j \geqslant \sum_{i,j} g((a_i \wedge b_j) + (b_i \wedge a_j))x_i x_j, \quad \forall x \in \Re^n, \qquad (GSI>)$$

holds whenever $g : J \to \Re$ is non-decreasing and convex.

Proof. By ($GSA<$), we know that, for all $n \geqslant 1$,

$$\sum_{1\leqslant i,j\leqslant n} f(a_i \vee a_j + b_i \vee b_j)x_i x_j \leqslant \sum_{1\leqslant i,j\leqslant n} f(a_i \vee b_j + b_i \vee a_j)x_i x_j, \quad \forall x \in \Re^n, \tag{5.39}$$

holds for every non-decreasing and concave function. If we replace in (5.39) the a_i and b_i with $-a_i$ and $-b_i$ we get

$$\sum_{1\leqslant i,j\leqslant n} f(-(a_i \wedge a_j + b_i \wedge b_j))x_i x_j \leqslant \sum_{1\leqslant i,j\leqslant n} f(-(a_i \wedge b_j + b_i \wedge a_j))x_i x_j, \quad \forall x \in \Re^n. \tag{5.40}$$

Let $g(x) = -f(-x)$. With this new function, (5.40) becomes

$$\sum_{1\leqslant i,j\leqslant n} g(a_i \wedge a_j + b_i \wedge b_j)x_i x_j \geqslant \sum_{1\leqslant i,j\leqslant n} g(a_i \wedge b_j + b_i \wedge a_j)x_i x_j. \tag{5.41}$$

Now let g be any non-decreasing convex function. Then $f(x) = -g(-x)$ is concave and non-decreasing. Thus for f, $GSA<$ holds, hence $GSI>$ holds for g. □

Theorem 5.12. *If g is non-decreasing and convex, then for all $n \geqslant 1$,*

$$\sum_{i,j} g((a_i \wedge a_j)(b_i \wedge b_j))x_i x_j \geqslant \sum_{i,j} g((a_i \wedge b_j)(b_i \wedge a_j))x_i x_j \quad \forall a, b \in [0, \infty)^n, \quad x \in \Re^n. \qquad (GPI>)$$

In particular, inequality $QPI>$ holds. Explicitly,

$$\sum_{i,j} \min(a_i, a_j) \min(b_i, b_j)x_i x_j \geqslant \sum_{i,j} \min(a_i, b_j) \min(a_j, b_i)x_i x_j, \quad \forall a, b \in [0, \infty)^n, x \in \Re^n, \tag{5.42}$$

hence $SPI>$ is true, too.

Proof. We may assume that all the a_i and b_i are positive. For if $1 \leqslant r \leqslant n$ is such that either $a_r = 0$ or $b_r = 0$, then, on each side of (5.42), each term with $i = r$ or $j = r$ equals zero. But then that index r may be dropped. Assuming that all the a_i and b_i are positive, we can write e^{s_i} instead of a_i and e^{t_i} instead of b_i. The inequality to be proved *GPI*> becomes

$$\sum_{i,j} g\big((e^{s_i} \wedge e^{s_j})(e^{t_i} \wedge e^{t_j})\big)x_i x_j \geqslant \sum_{i,j} g\big((e^{s_i} \wedge e^{t_j})(e^{t_i} \wedge e^{s_j})\big)x_i x_j \tag{5.43}$$

or

$$\sum_{i,j} g\big(e^{s_i \wedge s_j + t_i \wedge t_j}\big)x_i x_j \geqslant \sum_{i,j} g\big(e^{s_i \wedge t_j + s_j \wedge t_i}\big)x_i x_j. \tag{5.44}$$

Let $f(x) = g(e^x)$. Then f is also non-decreasing and convex. Thus (5.44) becomes

$$\sum_{i,j} f(s_i \wedge s_j + t_i \wedge t_j)x_i x_j \geqslant \sum_{i,j} f(s_i \wedge t_j + t_i \wedge s_j)x_i x_j \tag{5.45}$$

which is the inequality *GSI*> which we know is true according to Theorem 5.11. □

Remark. Why does *SPA*< not hold? Consider

$$\sum_{1 \leqslant i,j \leqslant n} g((a_i \vee a_j) \times (b_i \vee b_j))x_i x_j \leqslant \sum_{1 \leqslant i,j \leqslant n} g((a_i \vee b_j) \times (b_i \vee a_j))x_i x_j. \tag{$GPA<$}$$

As in the previous proof, write e^{s_i} instead of a_i and e^{t_i} instead of b_i. The inequality *GPA*< becomes

$$\sum_{i,j} g(e^{s_i \vee s_j + t_i \vee t_j}) \leqslant \sum_{i,j} g(e^{s_i \vee t_j + s_j \vee t_i}).$$

According to Theorem 5.7, this inequality holds if the function $f(x) = g(e^x)$ is concave and non-decreasing. But this inequality fails if $g(x) = x$ because then *GPA*< becomes *SPA*< which we know to be false. That may explain the failure of *SPA*<.

Remark. The previous example suggests that *GSA*< in Theorem 5.7 is false as soon as f fails to be increasing or if f is *not* concave, meaning that $S < 0$ (S as defined at (5.31)) for at least one choice of n, a, b, x. That suggestion is true.

Proposition 5.13. *If f is not non-decreasing, then GSA< is false.*

Proof. Let $p < q$ be such that $f(p) > f(q)$. Let $x = p - q/2$, $y = q/2$. Let $a_i = x$ and $b_i = y, \forall\, 1 \leqslant i \leqslant n$. Then property *GSA*< (if true) would say that $f(x + y) \leqslant f(2(x \vee y)) = f(2y) \Leftrightarrow f(p) \leqslant f(q)$, contradicting $f(p) > f(q)$. □

Counterexample. If f is increasing but *not* concave, then *GSA*< may fail.

Let $n = 2$, $a = (0, 1)$, $b = (2, -1)$. Thus $b_2 < a_1 < a_2 < b_1$ and $U = \begin{pmatrix} 2 & 3 \\ 3 & 0 \end{pmatrix}$, $V = \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix}$. Then (5.31) yields that $S = Ax_1^2 + Bx_2^2 + 2Cx_1x_2$ with $A = f(4) - f(2)$, $B = f(2) - f(0)$ and $C = f(2) - f(3)$. As $A, B \geqslant 0$, for this quadratic form to be non-negative we must have $\Delta \geqslant 0$ where $\Delta = AB - C^2$. If f is concave this is of course true, but it is easy to find continuous increasing functions for which $\Delta < 0$ so that *GSA*< cannot hold for them.

So far we have assumed that a_i and b_i are non-negative. That forced the matrices U, V to be non-negative, too. So any inequality of the form *QEF* implied the corresponding *REF*. But what happens if we allow $a, b \in \Re^n$ to be arbitrary? Sometimes *QEF* holds in this generalized form. But what about *REF*? Here is a counterexample.

Proposition 5.14. *If $a, b \in \Re^n$, then QSP>, QPS< and RSP> hold but RPS< is false.*

Proof. The proof of *QSP*> and *QPS*< reduces to the obvious inequality $p^2 + q^2 \geqslant 2pq$ which holds for any real numbers. The fact that *RPS*< is false can be seen by choosing $n = 2$, $a = (1, 2)$ and $b = (0, -1)$. Then $U = \begin{pmatrix} 0 & -3 \\ -3 & -8 \end{pmatrix}$, $V = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. The eigenvalues of U are 1 and -9 thus $\rho(U) = 9$. V has a double eigenvalue equal to 1 hence $\rho(V) = 1$. Thus $W = \begin{pmatrix} 1 & 3 \\ 3 & 9 \end{pmatrix}$ is positive semidefinite, but $\rho(U) > \rho(V)$, disproving *RPS*<.

However, *RSP*> is true. Indeed, the matrices U and V are given by

$$u_{i,j} = a_i a_j + b_i b_j, \quad v_{i,j} = a_i b_j + b_i a_j. \tag{5.46}$$

We shall prove that $\rho(U) \geqslant \rho(V)$.

Let $\lambda \neq 0$ be an eigenvalue of U. Therefore there exists an $x \neq 0$ such that $Ux = \lambda x$. But

$$\begin{aligned}(Ux)_i &= \sum_{j=1}^{n} u_{i,j} x_j = \sum_{j=1}^{n} (a_i a_j + b_i b_j) x_j \\ &= \sum_{j=1}^{n} a_i a_j x_j + \sum_{j=1}^{n} b_i b_j x_j = \alpha a_i + \beta b_i,\end{aligned}$$

where

$$\alpha = \alpha(x) = \langle a, x \rangle = \sum_{j=1}^{n} a_j x_j \quad \text{and} \quad \beta = \beta(x) = \langle b, x \rangle = \sum_{j=1}^{n} b_j x_j. \tag{5.47}$$

The equation $Ux = \lambda x$ becomes $\alpha a + \beta b = \lambda x$. Therefore, multiplying on the left by the row vector a, we get

$$\lambda\alpha = \sum_{j=1}^{n} a_j \lambda x_j = \sum_{j=1}^{n} a_j(\alpha a_j + \beta b_j) = A\alpha + C\beta,$$

where

$$A = \sum_{j=1}^{n} a_j^2, \quad B = \sum_{j=1}^{n} b_j^2 \quad \text{and} \quad C = \sum_{j=1}^{n} a_j b_j \quad (\text{thus } C^2 \leqslant AB). \tag{5.48}$$

Multiplying on the left by the row vector b, we get

$$\lambda\beta = \sum_{j=1}^{n} b_j \lambda x_j = \sum_{j=1}^{n} b_j(\alpha a_j + \beta b_j) = C\alpha + B\beta.$$

Thus λ must satisfy

$$\alpha(\lambda - A) - C\beta = 0, \quad C\alpha - \beta(\lambda - B) = 0. \tag{5.49}$$

If we think of (5.49) as a homogeneous system of linear equations with unknowns α and β, we want it to have non-trivial solutions (since $\alpha = \beta = 0 \Rightarrow \lambda x = 0 \Rightarrow \lambda = 0$, because x was supposed to be an eigenvector). The condition for the existence of non-trivial solutions is that

$$\det\begin{pmatrix} \lambda - A & -C \\ -C & \lambda - B \end{pmatrix} = 0 \iff \lambda^2 - \lambda(A + B) + AB - C^2 = 0. \tag{5.50}$$

So any non-zero eigenvalue of U must satisfy (5.50). Both roots of (5.50) being positive, the spectral radius of U is the greater of the two roots:

$$\rho(U) = \frac{1}{2}\left(A + B + \sqrt{(A - B)^2 + 4C^2}\right). \tag{5.51}$$

Now the eigenvalues λ of V should satisfy $Vx = \lambda x$ with some $x \neq 0$. Similar computations yield

$$\begin{aligned}(Vx)_i &= \sum_{j=1}^{n} v_{i,j} x_j = \sum_{j=1}^{n} (a_i b_j + b_i a_j) x_j \\ &= \sum_{j=1}^{n} a_i b_j x_j + \sum_{j=1}^{n} b_i a_j x_j = \beta a_i + \alpha b_i = \lambda x_j,\end{aligned}$$

hence

$$\lambda\beta = \sum_{j=1}^{n} b_j \lambda x_j = \sum_{j=1}^{n} b_j(\beta a_j + \alpha b_j) = C\beta + B\alpha,$$

$$\lambda\alpha = \sum_{j=1}^{n} a_j \lambda x_j = \sum_{j=1}^{n} a_j(\beta a_j + \alpha b_j) = A\beta + C\alpha.$$

Thus α and β satisfy

$$B\alpha - (\lambda - C)\beta = 0, (C - \lambda)A + A\beta = 0. \tag{5.52}$$

If non-degenerate solutions exist, we must have

$$\det \begin{pmatrix} B & C-\lambda \\ C-\lambda & A \end{pmatrix} = 0 \iff \lambda^2 - 2C\lambda + C^2 - AB = 0. \tag{5.53}$$

So $\rho(V)$ is the greater of the magnitudes of the two roots of (5.53):

$$\rho(V) = \max\left(|C - \sqrt{AB}|, |C + \sqrt{AB}|\right) = |C| + \sqrt{AB}. \tag{5.54}$$

It is easy to see that $\rho(V) \leqslant \rho(U)$. Indeed, let $\lambda = |C| + \sqrt{AB}$. So $\lambda \geqslant 0$ and $\lambda^2 = 2C\lambda + \Delta$, where $\Delta = AB - C^2 = \lambda(\sqrt{AB} - |C|) \geqslant 0$. We want to check that λ lies between the two roots of (5.50) or equivalently that $\lambda^2 - \lambda(A + B) + AB - C^2 \leqslant 0 \Leftrightarrow \lambda^2 - \lambda(A + B) + \lambda(\sqrt{AB} - |C|) \leqslant 0 \Leftrightarrow \lambda(\lambda - A - B + \sqrt{AB} - |C|) \leqslant 0 \Leftrightarrow -A - B + 2\sqrt{AB} \leqslant 0$ which is obvious. $\square$

5.1. *Generalizing the inequalities*

It is not important that the index set be finite.

A technique exists to generalize all the inequalities of the form *SEF*. Suppose for instance that *SEF*< holds, namely,

$$\sum_{1 \leqslant i,j \leqslant n} u_{i,j} \leqslant \sum_{1 \leqslant i,j \leqslant n} v_{i,j}, \tag{5.55}$$

where $u_{i,j} = E(F(a_i, a_j), F(b_i, b_j))$ and $v_{i,j} = E(F(a_i, b_j), F(b_i, a_j))$ and a, b are any vectors of length n, for any positive integer n.

If one replaces a and b with a^*, b^* constructed by repeating each pair (a_i, b_i) k_i times (k_i non-negative integers), then (5.55) becomes

$$\sum_{1 \leqslant i,j \leqslant n} u_{i,j}k_ik_j \leqslant \sum_{1 \leqslant i,j \leqslant n} v_{i,j}k_ik_j \quad \forall n, k_i \text{ non-negative integers.} \tag{5.56}$$

But (5.56) implies that a similar inequality must hold with k_i replaced by rational numbers p_i, $1 \leqslant i \leqslant n$. All these p_i can be written as k_i/k with the same k. So we have

$$\sum_{1 \leqslant i,j \leqslant n} u_{i,j}p_ip_j \leqslant \sum_{1 \leqslant i,j \leqslant n} v_{i,j}p_ip_j \quad \forall (p_i)_{1 \leqslant i \leqslant n} \text{ rational non-negative numbers.} \tag{5.57}$$

Now let $(p_i)_{1 \leqslant i \leqslant n}$ be any non-negative numbers. Consider sequences of non-negative rationals $p_{i,m}$ converging to p_i as $m \to \infty$. As (5.57) holds for $p_{i,m}$ instead of p_i for any m, it holds in the limit, too. Therefore

$$\sum_{1\leqslant i,j\leqslant n} E(F(a_i,a_j),F(b_i,b_j))p_i p_j \leqslant \sum_{1\leqslant i,j\leqslant n} E(F(a_i,b_j),F(b_i,a_j))p_i p_j \quad \forall (p_i)_{1\leqslant i\leqslant n} \geqslant 0. \tag{5.58}$$

Now consider a finite measure space (Ω, K, μ), a partition of Ω (*not* the set $\Omega = \{(p,q,r) \in [-\infty,\infty]^3;\ E_pE_qE_r<$ or $E_pE_qE_r>$ hold(s)$\}$ defined before Theorem 2.10), namely $(A_i)_{1\leqslant i\leqslant n}$ and two simple functions $f = \sum_{i=1}^n a_i 1_{A_i}$, $g = \sum_{i=1}^n b_i 1_{A_i}$. (Any pair of simple functions can be written that way.) Let $p_i = \mu(A_i)$. Then

$$\int\int E(F(f(x),f(y)),F(g(x),g(y)))\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) = \sum_{1\leqslant i,j\leqslant n} E(F(a_i,a_j),F(b_i,b_j))p_i p_j$$

and

$$\int\int E(F(f(x),g(y)),F(g(x),f(y)))\,\mathrm{d}\mu(x)\mathrm{d}\mu(y) = \sum_{1\leqslant i,j\leqslant n} E(F(a_i,b_j),F(b_i,a_j))p_i p_j.$$

So

$$\int\int E(F(f(x),f(y)),F(g(x),g(y)))\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) \leqslant \int\int E(F(f(x),g(y)),F(g(x),f(y)))\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)$$

must hold for simple f, g. Approximating measurable functions as usual by simple ones, we conclude:

Theorem 5.15. *If SEF< holds, then*

$$\int\int E(F(f(x),f(y)),F(g(x),g(y)))\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) \leqslant \int\int E(F(f(x),g(y)),F(g(x),f(y)))\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)$$

holds too, for any measurable non-negative functions f, g, where μ is a positive measure. A similar generalization holds for SEF>.

Corollary 5.16. *If SEF< holds and X, Y are two independent identically distributed random variables and f, g are two measurable non-negative functions, then*

$$\mathbf{E}(E(F(f(X), f(Y)), F(g(X), g(Y)))) \\ \leqslant \mathbf{E}(E(F(f(X), g(Y)), F(g(X), f(Y)))). \tag{5.59}$$

Proof. Apply the standard transport formula: Let μ be a measure on some space E. Let $f : E \to F$ be a measurable function, and on F define the image measure $\upsilon(B) = \mu(\{x; f(x) \in B\})$. If $g : F \to \Re$ is measurable, then $\int g \, \mathrm{d}\nu = \int g(f) \, \mathrm{d}\mu$. □

Remark. Similar reasoning shows that if *QEF*< holds, then (5.56) may be replaced by

$$\sum_{1 \leqslant i,j \leqslant n} u_{i,j} k_i k_j \leqslant \sum_{1 \leqslant i,j \leqslant n} v_{i,j} k_i k_j \quad \forall k \in \Re^n. \tag{5.60}$$

Then the usual approximation of measurable functions by simple ones gives:

Theorem 5.17. *If QEF< holds, then*

$$\int\int E(F(f(x), f(y)), F(g(x), g(y))) \, \mathrm{d}\mu(x) \, \mathrm{d}\mu(y) \\ \leqslant \int\int E(F(f(x), g(y)), F(g(x), f(y))) \, \mathrm{d}\mu(x) \, \mathrm{d}\mu(y),$$

for any measurable non-negative *functions f, g and any bounded signed measure μ. A similar generalization holds for SEF>.*

Whenever we can replace S with Q in an inequality *SEF*, then we can replace μ from Theorem 5.15 (a positive measure) with a signed measure. So far, we have no counterexamples to the Conjecture that opens this section.

6. Inequalities of the form *PEF*

Our proof will be roughly as follows.

PIP	equivalent to *SIS*	
PIS	implied by *PAS*	
PIA	equivalent to *SIA*	
PPI	from Corollary 4.10	
PPS	from Corollary 4.10	
PPA	from Corollary 4.10	
PSI	false (Theorem 6.7)	(6.1)
PSP	from Theorem 6.1	
PSA	from *GSA* (Corollary 5.9)	
PAI	equivalent to *SAI*	
PAP	equivalent to *SAS*	
PAS	from Proposition 6.4	

Theorem 6.1. *The inequality PSP>, namely,*

$$\prod_{1\leqslant i,j\leqslant n}(a_i a_j + b_i b_j) \geqslant \prod_{1\leqslant i,j\leqslant n}(a_i b_j + a_j b_i), \tag{6.2}$$

holds for all $n \geqslant 1$ *and all* $a, b \in [0, \infty)^n$.

Proof. Step 1. *Preliminaries*. By continuity, one may assume that $a_i > 0$ for all i. Letting $t_i = b_i/a_i$, (6.2) is equivalent to

$$\prod_{i,j}(1 + t_i t_j) \geqslant \prod_{i,j}(t_i + t_j) \quad \text{if } t_i \geqslant 0 \text{ for all } i. \tag{6.3}$$

This inequality is trivially true for $n = 1$. When $n = 2$,

$$\prod_{i,j}(1 + t_i t_j) - \prod_{i,j}(t_i + t_j) = (1 - t_1 t_2)^2[1 + t_1^2 + t_2^2 + 4t_1 t_2 + t_1^2 t_2^2] \geqslant 0. \tag{6.4}$$

The equality sign holds if and only if $t_1 t_2 = 1$. The function

$$f(t) = f_n(t) = f_n(t_1, \ldots, t_n) = \prod_{i,j}\frac{t_i + t_j}{1 + t_i t_j}, \quad \text{for } t \in \Re^n_+, \tag{6.5}$$

is non-negative, continuous and analytic everywhere on $\Re^n_+$. In addition, $f_n(t_1, \ldots, t_n)$ is symmetric, that is, invariant under all $n!$ permutations. Moreover, $f_n(t) = 0$ if and only if at least one of the coordinates t_j vanishes. Otherwise, $f_n(t) > 0$; and then the value $f_n(t)$ remains unchanged when (simultaneously) each coordinate t_j is replaced by its reciprocal. In view of (6.5), (6.2) is equivalent to

$$f_n(t) \leqslant 1 \quad \text{for all } n \geqslant 1 \text{ and each } t \in \Re^n_+. \tag{6.6}$$

Definition. A point $t = (t_1, \ldots, t_n) \in \Re^n_+$ has *elementary structure* if $\{t_1, \ldots, t_n\}$ completely decomposes into

$$\text{singlets } \{t_j = 1\} \text{ and pairs } (t_r, t_s) \text{ such that } 0 < t_r < 1 < t_s \text{ and } t_r t_s = 1. \tag{6.7}$$

For example, $n = 6$ and $t = (3, 1, 2, 1/3, 1, 1/2)$ has elementary structure. Our induction hypothesis will be

Property $E(n)$. Property $E(n)$ holds if (6.6) is true and, moreover, $f_n(t) = 1$ if and only if $t \in \Re^n_+$ has elementary structure.

Definition. A point $t \in \Re^n_+$ is special if either

$$t_j = 0 \text{ or } t_j = 1 \text{ for some } j; \quad \text{or else } t_r t_s = 1 \text{ for some } r, s \text{ with } r \neq s. \tag{6.8}$$

Here $j, r, s \in \{1, 2, \ldots, n\}$. All other points $t \in \Re_+^n$ are non-special. Thus $t \in \Re_+^n$ is non-special if and only if

$$t_j > 0; \quad t_r t_s \neq 1 \text{ for any } j, r, s, \in \{1, 2, \ldots, n\}. \tag{6.9}$$

The set of all non-special points $t \in \Re_+^n$ is an open subset of $(0, \infty)^n$. Almost all $t \in \Re_+^n$ are non-special.

Main Theorem. *Property $E(n)$ holds for all n.*

Proof. Proof by induction on n. Property $E(1)$ is trivially true, while (6.4) shows that $E(2)$ is true. From now on n is fixed with $n \geqslant 3$. We will show that $E(n)$ is true, assuming (from now on) that $E(m)$ is true for all $1 \leqslant m \leqslant n - 1$.

To do the induction, we prove some relations between f_n and f_{n+1}.

Step 2. If t is special and $E(m)$ is true for $1 \leqslant m \leqslant n - 1$, then $f_n(t) \leqslant 1$ and $f_n(t) = 1$ if and only if t has elementary structure.

Let $t \in [0, \infty)^n$ and $x, y \geqslant 0$. Then the reader is invited to check that

$$f_{n+1}(t_1, \ldots, t_n, x) = f_n(t)\rho(t, x), \tag{6.10}$$

where

$$\rho(t, x) = \frac{2x}{1 + x^2} \left(\prod_{j=1}^{n} \frac{t_j + x}{1 + t_j x} \right)^2 \tag{6.11}$$

and

$$f_{n+2}(t_1, \ldots, t_n, x, y) = f_n(t)\sigma(t, x, y), \tag{6.12}$$

where

$$\sigma(t, x, y) = \frac{2x}{1 + x^2} \frac{2y}{1 + y^2} \left(\frac{x + y}{1 + xy} \prod_{j=1}^{n} \frac{t_j + x}{1 + t_j x} \frac{t_j + y}{1 + t_j y} \right)^2. \tag{6.13}$$

Remark that

$$\rho(t, 0) = 0, \quad \rho(t, 1) = 1 \quad \text{and} \quad \sigma(t, x, 1/x) = 1. \tag{6.14}$$

Let t be a special point. Then exactly one of three cases holds:

1. For some j, $t_j = 0$. If $t^* \in [0, \infty)^{n-1}$ is the vector obtained from t after deleting the component t_j, then $f_n(t) = f_{n-1}(t^*)\rho(t^*, 0) = 0$ by (6.10).
2. For some j, $t_j = 1$. Then $f_n(t) = f_{n-1}(t^*)\rho(t^*, 1) = f_{n-1}(t^*)$ by (6.10), where $t^* \in [0, \infty)^{n-1}$ is the vector obtained from t after deleting the component t_j.
3. There exist r, s with $r \neq s$ such that $t_r t_s = 1$. Then $f_n(t) = f_{n-2}(t^*)\sigma(t^*, t_r, t_s) = f_{n-2}(t^*)\sigma(t^*, t_r, 1/t_r) = f_{n-2}(t^*)$ by (6.10), where $t^* \in [0, \infty)^{n-2}$ is the vector obtained from t after deleting the components t_r and t_s.

Assuming that $E(m)$ holds for $m \leqslant n-1$, in all cases $f_n(t) < 1$.

Now suppose that $f_n(t) = 1$. That can happen only in the last two cases. In Case 2, $f_{n-1}(t^*) = 1$, hence t^* has elementary structure. Inserting a component equal to 1 somewhere in t does not affect its elementary structure. In Case 3, $f_{n-2}(t^*) = 1$, thus t^* has elementary structure. Inserting a pair of components $(x, 1/x)$ somewhere in t does not change its elementary structure. □

Remark. If $t \in \Re_+^n$ has elementary structure, then $f_n(t) = 1$, as follows from an easy induction on n.

Step 3. If $(\operatorname{grad} f)(t) = 0$ and $t \in (0, \infty)^n$, then t must be special.

Let $g_n(t) = \log f_n(t)$. Taking into account the part of $g_n(t)$ involving the coordinate t_r, one easily sees that

$$\begin{aligned}\frac{\partial}{\partial t_r} g_n(t) &= \frac{1}{t_r} - \frac{2t_r}{1+t_r^2} + 2\sum_{j \neq r}\left\{\frac{1}{t_r + t_j} - \frac{t_j}{1+t_r t_j}\right\} \\ &= 2\sum_{j=1}^{n}\left\{\frac{1}{t_r + t_j} - \frac{t_j}{1+t_r t_j}\right\} = 2\sum_{j=1}^{n}\frac{1-t_j^2}{(t_r+t_j)(1+t_r t_j)}. \end{aligned} \tag{6.15}$$

Therefore

$$\begin{aligned}(\operatorname{grad} f)(t) = \mathbf{0} &\iff (\operatorname{grad} g)(t) = \mathbf{0} \\ &\iff \sum_{j=1}^{n}\frac{1-t_j^2}{(t_r+t_j)(1+t_r t_j)} = 0, \quad \forall 1 \leqslant r \leqslant n.\end{aligned}$$

Consider the function

$$h_n(t, z) = 2\sum_{j=1}^{n}\left\{\frac{1}{t_j + z} - \frac{t_j}{1+t_j z}\right\} = 2\sum_{j=1}^{n}\frac{1-t_j^2}{(z+t_j)(1+z t_j)}. \tag{6.16}$$

By the change of variables

$$w = \frac{1}{2}\left(z + \frac{1}{z}\right), \quad w_j = w_j(t) = \frac{1}{2}\left(t_j + \frac{1}{t_j}\right) \tag{6.17}$$

(so that $z = t_j$ corresponds to $w = w_j$), (6.16) simplifies to

$$h_n(t, z) = \frac{1}{z} H_n(t, w), \quad \text{where } H_n(t, w) = \sum_{j=1}^{n}\frac{\frac{1}{t_j} - t_j}{w + w_j}. \tag{6.18}$$

Then (6.15) can be written as

$$\frac{\partial}{\partial t_r} g_n(t) = \frac{1}{t_r} H_n(t, w_r) \quad \text{for } r = 1, \ldots, n. \tag{6.19}$$

Now suppose that $t^0 \in (0, \infty)^n$ is a stationary point, that is, $\mathrm{grad}(f)(t^0) = 0$. By (6.19), that would imply that

$$H_n(t^0, w_j) = 0, \quad \forall 1 \leqslant j \leqslant n, \quad \text{where } w_j = \frac{1}{2}\left(\frac{1}{t_j^0} + t_j^0\right). \tag{6.20}$$

We shall prove Step 3 by contradiction. Suppose *ad absurdum* that t^0 is non-special and (6.20) holds. That is, for no r do we have that $t_r^0 = 1$ (that is, $w_r = 1$) nor does it happen that $t_r^0 t_s^0 = 1$ when $r \neq s$. Hence, if $r \neq s$, then $w_r = w_s$ if and only if $t_r^0 = t_s^0$ (since if $x + 1/x = y + 1/y$, then either $x = y$ or $y = 1/x$). In other words, the number of different t_j is the same as the number of different w_j. Let then $1 < \omega_1 < \omega_2 < \cdots < \omega_k$, for $1 \leqslant k \leqslant n$, be the ordered set of different values among $w_1, \ldots, w_n$. Let further

$$J_q = \{j; w_j = \omega_q, j = 1, \ldots, n\} \quad \text{for } q = 1, \ldots, k.$$

All the values t_j with $j \in J_q$ are equal to one and the same value τ_q such that $\tau_q > 0$, and $\tau_q \neq 1$ and $\frac{1}{2}\left(t_q + \frac{1}{t_q}\right) = \omega_q$. Let $n_q = |J_q|$. Thus $n_q \geqslant 1$; $n_1 + \cdots + n_k = n$. It follows from (6.17) that

$$H_n(t^0, w) = \sum_{q=1}^{k} n_q \frac{\frac{1}{\tau_q} - \tau_q}{w + \omega_q}. \tag{6.21}$$

From (6.19), $H_n(t^0, \omega_q) = 0$ for $q = 1, \ldots, k$. Now consider the polynomial

$$\varphi(w) = H_n(t^0, w)\prod_{q=1}^{k}(w + w_q) = \sum_{q=1}^{k} n_q \left(\frac{1}{\tau_q} - \tau_q\right)\prod_{p \neq q}(w + \omega_p). \tag{6.22}$$

This polynomial $\varphi(w)$ is of degree at most $k - 1$, while from (6.20), $\varphi(w)$ vanishes at the k distinct points $\omega_1, \ldots, \omega_k$. This is possible only when $\varphi(w) = 0$ and thus $H_n(t^0, w) = 0$. But $H_n(t^0, w) = 0$ is false because, for $q = 1, \ldots, k$, the meromorphic function $w \to H_n(t^0, w)$ has at $w = -\omega_q$ the residue $n_q(1/\tau_q - \tau_q)$, which is non-zero (as $\tau_q \neq 1$ and t is non-special). Thus the assumption that t^0 is a stationary non-special point leads to a contradiction. This proves Step 3.

Corollary

(i) *If f_n has a local maximum at $t^0 \in (0, \infty)^n$, then t^0 must be a special point, hence $f_n(t^0) \leqslant 1$ by Step* 2.

(ii) *Let K be any non-empty subset of $\Re_+^n$ such that* $\sup\{f_n(t); t \in K\}$ *is assumed at a point $t^0 \in \mathrm{int}(K)$. Then t^0 is special and* $\sup\{f_n(t); t \in K\} = f_n(t^0) \leqslant 1$.

(iii) *If* $\sup f_n = \sup\{f_n(t); t \in \Re_+^n\}$ *is attained at some point $t^0 \in \Re_+^n$, then Property $E(n)$ is true. That is, if the supremum is attained, then it must be attained at a special point and thus it is equal to* 1.

Step 4. The supremum is attained.

Let $1 < c < \infty$ and $K(c)$ be the compact cube

$$K(c) = \left\{t \in \Re_+^n; 1/c \leqslant t_j \leqslant c \ \text{ for } j = 1, \ldots, n\right\}. \tag{6.23}$$

Let further

$$M(c) = \max\{f_n(t); t \in K(c)\}. \tag{6.24}$$

In proving Property E(n), it suffices to show that

$$M(c) \leqslant 1 \quad \text{for all } c > 1. \tag{6.25}$$

Namely, each point $t \in \Re^n_+$ with $f_n(t) > 0$ (that is, $t_j > 0$ for all j) is contained in $K(c)$ as soon as c is sufficiently large. Hence, (6.25) would imply that $\sup f_n = 1$. In addition, if t is such that $f_n(t) = 1$, then t would clearly be stationary and thus special and thus of elementary structure (see Steps 2 and 3).

To prove (6.25), we will derive a contradiction from the assumption that, for some fixed c, $c > 1$,

$$M(c) > 1. \tag{6.26}$$

Since $K(c)$ is compact, there exists $t^0 \in K(c)$ (to be kept fixed) such that

$$f_n(t^0) = M(c) > 1. \tag{6.27}$$

It follows from Step 2 that the point t^0 must be non-special. Moreover, from Step 3 it is impossible that $t^0 \in \operatorname{int}(K(c))$. That is, t^0 must be a boundary point of $K(c)$. Thus the coordinates t^0_j of t^0 satisfy $1/c \leqslant t^0_j \leqslant c$, for all j, while either $t^0_j = c$ or $t^0_j = 1/c$ for at least one index j. Replacing each t^0_j by its reciprocal, if necessary, we may as well assume that $t^0_j = c$ for some j. Let $w_j = (1/t^0_j + t^0_j)/2$. Thus $1 < w_j \leqslant (1/c + c)/2$. Since t^0 is non-special, we know that $w_r = w_s$ if and only if $t^0_r = t^0_s$.

The following machinery was previously used in the proof of Step 3. Let $1 < \omega_1 < \omega_2 < \cdots < \omega_k$, for $1 \leqslant k \leqslant n$, be the ordered set of different values among $w_1, \ldots, w_n$. In particular, $\omega_k = (1/c + c)/2$. Further $k \geqslant 2$. For if $k = 1$, then $t^0_j = c$ for $j = 1, \ldots, n$, thus $f_n(t^0) = \left(\frac{2c}{1+c^2}\right)^{n^2} < 1$, which contradicts (6.27). As before, let

$$J_q = \{j; w_j = \omega_q, j = 1, \ldots, n\}; \quad n_q = |J_q|, \;\; q = 1, \ldots, k. \tag{6.28}$$

Since t^0 is non-special, all the values t^0_j with $j \in J_q$ are equal to one and the same value $\tau_q > 0$, which is such that $(\tau_q + 1/\tau_q)/2 = \omega_q$. We already showed (see Step 3) that

$$\frac{\partial}{\partial t_r} g_n(t^0) = \frac{1}{t^0_r} H_n(t^0, w_r), \quad \text{for } r = 1, \ldots, n. \tag{6.29}$$

Here $g_n(t) = \log f_n(t)$ and

$$H_n(t^0, w) = \sum_{j=1}^{n} \frac{\frac{1}{t^0_j} - t^0_j}{w + w_j} = \sum_{q=1}^{k} \frac{R(q)}{w + \omega_q} \quad \text{where } R(q) = n_q \left(\frac{1}{\tau_q} - \tau_q\right). \tag{6.30}$$

$H_n(t^0, w)$ is a meromorphic function having $-\omega_q$ as a simple pole (with residue $R(q) \neq 0$), for $q = 1, \ldots, k$.

The (non-special) point $t^0 \in \partial K(c)$ has coordinates $t_j^0 = c$ when $j \in J_k$ while otherwise $1/c < t_j^0 < c$. Thus t^0 is located in the face F of the cube $K(c)$ defined by

$$F = \left\{t;\, 1/c < t_j < c \text{ if } j \notin J_k;\, t_j = c \text{ if } j \in J_k\right\}.$$

This face F is relatively open, that is, open relative to its affine span. Recall that $g_n(t)$ restricted to $K(c)$ takes its largest value at $t^0 \in F$. Hence, $g_n(t)$ restricted to the relatively open set F is also maximal at $t = t^0$, implying that

$$\frac{\partial}{\partial t_r} g_n(t^0) = 0 \quad \text{if } j \notin J_k, \quad j = 1, \ldots, n. \tag{6.31}$$

To arrive at a contradiction, it suffices to show that

$$\frac{\partial}{\partial t_r} g_n(t^0) < 0 \quad \text{if } j \in J_k. \tag{6.32}$$

(These $n_k = |J_k| \geqslant 1$ derivatives are all equal.) For if (6.32) were true, then, slightly moving away from the point $t^0 \in \partial K(c)$ into the interior of $K(c)$, by replacing each coordinate $t_j = c$ by a slightly smaller number ($j \in J_k$), one would encounter values $g_n(t)$ with $t \in \operatorname{int}(K(c))$ that are strictly larger than the starting value $g_n(t^0)$. This would contradict the assumed maximality of t^0.

In view of (6.29), with $r \in J_k$ and thus $t_r^0 = c > 0$, the desired result (6.32) is equivalent to

$$H_n(t^0, \omega_k) < 0, \omega_k = \frac{1}{2}\left(\frac{1}{c} + c\right). \tag{6.33}$$

From (6.29) and (6.31) we know that

$$H_n(t^0, \omega_q) = 0, \quad q = 1, \ldots, k-1, \quad 1 < \omega_1 < \omega_2 < \cdots < \omega_k. \tag{6.34}$$

Multiplying (6.30) by $\prod_{q=1}^{k}(w + w_q)$, one obtains the polynomial

$$\varphi(w) = \sum_{q=1}^{k} R(q) \prod_{s \neq q} (w + w_s) = C \prod_{q=1}^{k-1} (w - w_q) \quad \text{where } C = \sum_{q=1}^{k} R(q).$$

Thus $\varphi(w)$ must be precisely of degree $k - 1$, in particular $C \neq 0$. Moreover,

$$H_n(t^0, w) \approx \frac{C}{w} \quad \text{when } |w| \to \infty \text{ and } C \neq 0. \tag{6.35}$$

It is also clear that the meromorphic function $w \to H_n(t^0, w)$ cannot have any zeros besides the zeros $\omega_1, \ldots, \omega_{k-1}$, which themselves must be simple zeros. Consequently, $H_n(t^0, w)$ is of constant sign on $(-\infty, -\omega_k)$ and also of constant sign on $(\omega_{k-1}, +\infty)$. From (6.35), these two signs must be opposite. Since $R(k) = n_k(1/c - c) < 0$, with $R(k)$ as the residue at $w = -\omega_k$, one has $H_n(t^0, w) > 0$ if w is slightly smaller than $-\omega_k$, hence, also throughout $(-\infty, -\omega_k)$. Therefore, $H_n(t^0, w) < 0$ throughout (ω_{k-1}, ∞). In particular (6.33) is true. $\square$

Remark. There are other ways of ending the proof. For instance, since $H_n(t^0, w)$ has no zeros in the interval $(-\omega_{q+1}, -\omega_q)$, it follows that the (non-zero) residues $R(q)$ and $R(q+1)$ must be of opposite sign, for $q = 1, \ldots, k-1$. Since $R(\omega_k) = n_k(1/c - c) < 0$, it follows that $R(q)$ must be of sign $(-1)^{k+1-q}$ $(q = 1, \ldots, k)$. In particular, the residue $R(1)$ at $-\omega_1$ is of sign $(-1)^k$. Thus $H_n(t^0, w)$ has sign $(-1)^k$ in the interval $(-\omega_1, \omega_1)$. Hence, in the interval (ω_1, ω_2) the sign of $H_n(t^0, w)$ equals $(-1)^{k-1}$ (since it has ω_1 as a simple zero). And so on. Thus, for the interval (ω_{q-1}, ω_q), we find the sign $(-1)^{k+1-q}$, $q = 1, \ldots, k-1$, which is positive when $q = k-1$, that is, for the interval $(\omega_{k-2}, \omega_{k-1})$. Because of the further sign change at the point ω_{k-1}, we find that $H_n(t^0, w)$ must be negative throughout the interval $(\omega_{k-1}, +\infty)$. In particular, $H_n(t^0, \omega_k) < 0$, which is precisely (6.33). □

Next we prove that $E_pAS>$ is true for $p \leqslant 1$, which implies $PAS>$ (letting $p \to 0$) and provides other insights, too. The technology is the same as that used in Section 4.

Let $p > 0$ and $x, y \in (0, \infty)^n$. We use the notation x and y here because we will also use a and b. Let

$$u_{i,j} = (x_i + x_j) \vee (y_i + y_j),\ v_{i,j} = (x_i + y_j) \vee (y_i + x_j). \tag{6.36}$$

The claim is that

$$\left(\sum_{1\leqslant i,j\leqslant n} u_{i,j}^p\right)^{\frac{1}{p}} \geqslant \left(\sum_{1\leqslant i,j\leqslant n} v_{i,j}^p\right)^{\frac{1}{p}} \tag{$E_pAS>$}$$

which, because $p > 0$, is the same as

$$\sum_{1\leqslant i,j\leqslant n} u_{i,j}^p \geqslant \sum_{1\leqslant i,j\leqslant n} v_{i,j}^p \tag{6.37}$$

or

$$\sum_{1\leqslant i,j\leqslant n} ((x_i + x_j) \vee (y_i + y_j))^p \geqslant \sum_{1\leqslant i,j\leqslant n} ((x_i + y_j) \vee (y_i + x_j))^p. \tag{6.38}$$

Define $D_f : (0, \infty)^n \times (0, \infty)^n \to \Re$ by

$$D_f(x, y) = \sum_{1\leqslant i,j\leqslant n} (f((x_i + x_j) \vee (y_i + y_j)) - f((x_i + y_j) \vee (y_i + x_j))). \tag{6.39}$$

The task is to prove that if $f(u) = u^p$, $0 \leqslant p \leqslant 1$, then $D_f \geqslant 0$.

Let $\mathbf{C}_2$ be the set of all functions $f : (0, \infty) \to \Re$ such that $D_f \leqslant 0$. We know that the cone $\mathbf{C}_2$ has property (A) from Proposition 8.4. Then the functions f defined by $f(u) = -u^p$ belong to $\mathbf{C}_2$, for any p in $0 \leqslant p \leqslant 1$.

Proposition 6.2. *$E_pAS>$ holds for any $p < 0$.*

Proof. Let $f(x) = x^p$ for some $p < 0$. Then f belongs to C_2 by Corollary 8.2. Therefore

$$\sum_{1 \leqslant i,j \leqslant n} ((x_i + x_j) \vee (y_i + y_j))^p \leqslant \sum_{1 \leqslant i,j \leqslant n} ((x_i + y_j) \vee (y_i + x_j))^p.$$

If we raise that to the negative power $1/p$ we get $E_pAS>$ in this case. □

Proposition 6.3. *If* $0 < p \leqslant 1$, *then* $E_pAS>$ *holds.*

Proof. This property is equivalent to the fact that $u \mapsto -u^p$ belongs to C_2, which is stated by Corollary 8.2. □

Proposition 6.4. *PAS> is true.*

Proof. The cone C_2 contains the function $x \mapsto -\log x$, which means that

$$\sum_{1 \leqslant i,j \leqslant n} (\log((x_i + x_j) \vee (y_i + y_j)) - \log((x_i + y_j) \vee (y_i + x_j))) \geqslant 0.$$

That is *PAS>*. □

Remark. As *SAS>* is $E_1AS>$, the method of Proposition 6.4 gives another proof of *SAS>* different from that of Theorem 5.5.

Corollary 6.5
(i) *If* $p \in [-\infty, 1]$, *then* $E_pAS>$ *holds.*
(ii) $SAE_r>$ *holds if* $r \in [1, \infty]$ *and* $SIE_r<$ *holds if* $r \in [-\infty, 0]$.

Proof. (i) Propositions 6.2–6.4. (ii) According to Theorem 2.9, $E_pE_qE_r$ and $E_{tp}E_{tq}E_{tr}$ are equivalent whenever $t \neq 0$. So, if $0 < p \leqslant 1$, then $E_pAS>$ is equivalent to $SAE_{1/p} = SAE_r>$ with $r \geqslant 1$. If $p < 0$, then $E_pAS>$ is equivalent to $SIE_{1/p}< = SIE_r<$ for $r < 0$. □

Remark. We now see that the collection of points (p, q, r) belonging to Ω contains all the points of the form $(p, \pm\infty, 1)$ with $-1 \leqslant p \leqslant 1$. Considering Theorem 2.9, Ω also contains the points $(1, \infty, r)$ with $r \geqslant 1$ and $(1, -\infty, r)$ with $r \leqslant 0$.

Corollary. *Let* X, Y *be two independent and identically distributed non-negative random variables. Let* f, g *be measurable functions. If* $0 < p \leqslant 1$, *then*

$$\begin{aligned} &\mathbf{E}\big[((f(X)^p + f(Y)^p)^{1/p} \vee ((g(X)^p + g(Y)^p)^{1/p}\big] \\ &\quad \leqslant \mathbf{E}\big[((f(X)^p + g(Y)^p)^{1/p} + ((g(X)^p + f(Y)^p)^{1/p}\big]. \end{aligned} \tag{6.40}$$

Proof. Corollary 5.16. □

Theorem 6.6. *If* $0 < p \leqslant 1$ *or if* $p < 0$, *then* $E_pIS<$ *is true. Consequently*, $PIS<$ *is true*, *too.*

Proof. We know now that if $p \geqslant 1$ or $p \leqslant 0$, then the inequality $E_pAS>$ is true. Using the notations in (2.45)–(2.56), $E_pAS>$ states that

$$L_{p,\infty,1}(a,b) \geqslant R_{p,\infty,1}(a,b), \quad \forall a,b \in (0,\infty)^n. \tag{6.41}$$

There are two cases.

Case 1. $p \geqslant 1$. Raising (6.41) to the positive power p gives

$$(L_{p,\infty,1}(a,b))^p \geqslant (R_{p,\infty,1}(a,b))^p, \quad \forall a,b \in (0,\infty)^n. \tag{6.42}$$

We want to prove that $(L_{p,-\infty,1}(a,b))^p \leqslant (R_{p,-\infty,1}(a,b))^p$, $\forall a,b \in (0,\infty)^n$. Remark that

$$\begin{aligned}
&(L_{p,\infty,1}(a,b))^p + (L_{p,-\infty,1}(a,b))^p \\
&\quad = \sum_{1\leqslant i,j\leqslant n} ((a_i+a_j)\vee(b_i+b_j))^p + \sum_{1\leqslant i,j\leqslant n} ((a_i+a_j)\wedge(b_i+b_j))^p \\
&\quad = \sum_{1\leqslant i,j\leqslant n} ((a_i+a_j)^p + (b_i+b_j))^p \\
&\quad = \sum_{1\leqslant i,j\leqslant n} \left(x_i^{\frac{1}{p}} + x_j^{\frac{1}{p}}\right)^p + \left(y_i^{\frac{1}{p}} + y_j^{\frac{1}{p}}\right)^p = L_{1,1,r}(x,y),
\end{aligned}$$

where $x_i = a_i^p$, $y_i = b_i^p$ with $r = 1/p \geqslant 1$. Theorem 4.1 says that for $r \geqslant 1$, $E_1E_1E_r<$ holds. Therefore $L_{1,1,r}(x,y) \leqslant R_{1,1,r}(x,y)$. It follows that $(L_{p,\infty,1}(a,b))^p + (L_{p,-\infty,1}(a,b))^p \leqslant (R_{p,\infty,1}(a,b))^p + (R_{p,-\infty,1}(a,b))^p$ which, together with (6.42), implies that

$$(L_{p,-\infty,1}(a,b))^p \leqslant (R_{p,-\infty,1}(a,b))^p, \quad \forall a,b \in (0,\infty)^n.$$

As $p > 0$, that is the same as $L_{p,-\infty,1}(a,b) \leqslant R_{p,-\infty,1}(a,b)$, $\forall a,b \in (0,\infty)^n$.

Case 2. $p < 0$. Raising (6.41) to the negative power p, it becomes

$$(L_{p,\infty,1}(a,b))^p \leqslant (R_{p,\infty,1}(a,b))^p, \quad \forall a,b \in (0,\infty)^n. \tag{6.43}$$

We want to prove that $(L_{p,-\infty,1}(a,b))^p \geqslant (R_{p,-\infty,1}(a,b))^p$, $\forall a,b \in (0,\infty)^n$. The equality $(L_{p,\infty,1}(a,b))^p + (L_{p,-\infty,1}(a,b))^p = L_{1,1,r}(x,y)$ (with $r = 1/p < 0$) holds in this case, too. We know that if $r < 0$, then $E_1E_1E_r>$ holds (by combining Theorems 4.2 and 2.9). Therefore $L_{1,1,r}(x,y) \geqslant R_{1,1,r}(x,y)$. It follows that

$$\begin{aligned}
&(L_{p,-\infty,1}(a,b))^p + (L_{p,-\infty,1}(a,b))^p \\
&\quad \geqslant (R_{p,-\infty,1}(a,b))^p + (R_{p,-\infty,1}(a,b))^p
\end{aligned} \tag{6.44}$$

which, together with (6.43), implies that $(L_{p,\infty,1}(a,b))^p \geqslant (R_{p,\infty,1}(a,b))^p$, $\forall a,b \in (0,\infty)^n$. As $p < 0$, that is the same as $L_{p,\infty,1}(a,b) \leqslant R_{p,\infty,1}(a,b)$, $\forall a,b \in (0,\infty)^n$.

The inequality $PIS<$ is a limiting case of $E_pIS<$ when $p \to 0$. $\square$

Theorem 6.7. *The property $E_\alpha SI = E_\alpha E_1 E_{-\infty}$ is false for all $\alpha < 1$, and in particular for $\alpha = 0$ (i.e., PSI> is false).*

Proof. Choose $n = 3$, $a = (3, 2, 1)$, $b = (2, 1, 3)$. Let $U_{i,j} = a_i \wedge a_j + b_i \wedge b_j$, $V_{i,j} = a_i \wedge b_j + b_i \wedge a_j$. Then

$$(U_{i,j}) = \begin{bmatrix} 5 & 3 & 3 \\ 3 & 3 & 2 \\ 3 & 2 & 4 \end{bmatrix}, \quad (V_{i,j}) = \begin{bmatrix} 4 & 3 & 4 \\ 3 & 2 & 3 \\ 4 & 3 & 2 \end{bmatrix}.$$

It suffices to show that

$$\left\{ \sum_{i,j} [U_{i,j}]^\alpha \right\}^{1/\alpha} < \left\{ \sum_{i,j} [V_{i,j}]^\alpha \right\}^{1/\alpha}, \quad \text{for all } \alpha < 1; \quad \alpha \neq 0. \tag{6.45}$$

For $\alpha = 0$, (6.45) becomes *PSI*. Numerically, $\prod U_{i,j} = 19{,}440 < 20{,}736 = \prod V_{i,j}$, the opposite of what *PSI* would say. From now on, assume $\alpha < 1; \alpha \neq 0$. Then (6.45) is equivalent to $[2(2^\alpha) + 5(3^\alpha) + 4^\alpha + 5^\alpha]^{1/\alpha} < [2(2^\alpha) + 4(3^\alpha) + 3(4^\alpha)]^{1/\alpha}$. Writing $h(\alpha) = 3^\alpha - 2(4^\alpha) + 5^\alpha$, the last inequality is the same as $h(\alpha) < 0$ if $0 < \alpha < 1$ and as $h(\alpha) > 0$ if $\alpha < 0$. Writing $\phi(\alpha) = h(\alpha)/4^\alpha = -2 + (3/4)^\alpha + (5/4)^\alpha$, the last inequalities are the same as $\phi(\alpha) < 0$ for $0 < \alpha < 1$ and $\phi(\alpha) > 0$ for $\alpha < 0$. These inequalities follow from $\phi(0) = \phi(1) = 0$ and the fact that $\phi(\alpha)$ is everywhere strictly convex. $\square$

Remark. Most pairs $a, b \in [0, \infty)^n$ do satisfy the inequality (6.45) for all $\alpha > 1$; for instance, take $a = b$. Our counterexample is a somewhat exceptional pair. Another exceptional pair is given by $n = 3$, $a = (4, 0, 3)$ and $b = (0, 3, 4)$.

Theorem 6.8. *$QAE_r>$ holds if $r \in [1, \infty]$ and $QIE_r<$ holds if $r \in [-\infty, 0]$.*

Proof. Replace (6.39) with

$$D_f(x, y, \xi) = \sum_{1 \leqslant i,j \leqslant n} \big(f((x_i + x_j) \vee (y_i + y_j)) \\ - f((x_i + y_j) \vee (y_i + x_j)) \big) \xi_i \xi_j,$$

where $\xi \in \Re^n$ and let the definition of $\mathbf{C}_2$ and f be the same. Then Proposition 8.4 asserts that $\mathbf{C}_2$ has property (A) and thus contains our useful functions $-\log x$ and $-x^p, 0 < p \leqslant 1$. $\square$

7. Generalizations and counterexamples; review of open questions

7.1. The set Ω of triplets $(p, q, r) \in \Re^3$ such that $E_p E_q E_r$ is true

It is probably difficult to obtain an explicit exact description of $\Omega = \{(p, q, r) \in \Re^3; E_p E_q E_r \text{ is true}\}$. Numerical exploration might give a rough approximation of it.

We can get some precise information about Ω from several special cases of some previous generalized inequalities.

First, if $f(x) = x^{\alpha}$, $0 < \alpha \leqslant 1$, then f is an increasing concave function on $J = (0, \infty)$ and *GSA*< (Theorem 5.7) becomes

$$\sum_{i,j}[\max(a_i, a_j) + \max(b_i, b_j)]^{\alpha} \leqslant \sum_{i,j}[\max(a_i, b_j) + \max(a_j, b_i)]^{\alpha}, \quad (7.1)$$

whenever all $a_i > 0, b_j > 0$. Equivalently, since $0 < \alpha \leqslant 1$,

$$\left\{\sum_{i,j}[\max(a_i, a_j) + \max(b_i, b_j)]^{\alpha}\right\}^{1/\alpha} \leqslant \left\{\sum_{i,j}[\max(a_i, b_j) + \max(a_j, b_i)]^{\alpha}\right\}^{1/\alpha} \quad (7.2)$$

which asserts $E_{\alpha}SA<$, equivalently $E_{\alpha}E_{1}E_{\infty}<$, equivalently $(p, q, r) = (\alpha, 1, \infty) \in \Omega$, for all $0 < \alpha \leqslant 1$. Moreover, $PSA< = E_0SA$ is the limiting case when $\alpha \downarrow 0$ of $E_{\alpha}SA<$. Hence $(\alpha, 1, \infty) \in \Omega$, for all $0 \leqslant \alpha \leqslant 1$.

Next, choose $J = (0, \infty)$ and $f(x) = -x^{-\beta}$ where $\beta > 0$. Thus f is increasing and concave function on J. Then *GSA*< becomes

$$\sum_{i,j}[\max(a_i, a_j) + \max(b_i, b_j)]^{-\beta} \geqslant \sum_{i,j}[\max(a_i, b_j) + \max(a_j, b_i)]^{-\beta}, \quad (7.3)$$

for $a_i > 0, b_j > 0$ and $\beta > 0$, or equivalently

$$\left\{\sum_{i,j}[\max(a_i, a_j) + \max(b_i, b_j)]^{-\beta}\right\}^{-1/\beta} \leqslant \left\{\sum_{i,j}[\max(a_i, b_j) + \max(a_j, b_i)]^{-\beta}\right\}^{-1/\beta}. \quad (7.4)$$

Now (7.4) is the inequality $E_{-\beta}SA<$ equivalently $E_{-\beta}E_{1}E_{\infty}<$ equivalently $(-\beta, 1, \infty) \in \Omega$, for all $\beta > 0$. The limiting case $-\beta \downarrow -\infty$ is *ISA*<, i.e., $(-\infty, 1, \infty) \in \Omega$.

Next choose $J = (0, \infty)$ and $g(x) = x^{\alpha}$ with $\alpha \geqslant 1$. Since $g(x)$ is increasing and convex on J, it follows from *GSI*> (Theorem 5.11) that

$$\sum_{i,j}[\min(a_i, a_j) + \min(b_i, b_j)]^{\alpha} \geqslant \sum_{i,j}[\min(a_i, b_j) + \min(a_j, b_i)]^{\alpha}, \quad (7.5)$$

if all $a_i > 0, b_j > 0$ and $\alpha \geqslant 1$. Equivalently, $E_{\alpha}SI = E_{\alpha}E_1E_{-\infty}$ holds for all $\alpha \geqslant 1$.

Proposition 7.1

(i) *If* $-\infty \leqslant \alpha \leqslant 1$, *then* $(\alpha, 1, \infty) \in \Omega$, *that is*, $E_\alpha E_1 E_\infty$ *holds.*
(ii) *If* $\alpha \geqslant 1$ *then* $(\alpha, 1, -\infty) \in \Omega$, *that is*, $E_\alpha E_1 E_{-\infty}$ *holds.*
(iii) *If* $p \leqslant r$ *and* $r \geqslant 0$, *then* $(p, p, r) \in \Omega$.
(iv) *If* m *is a positive integer*, *then* $(m, 1, 0) \in \Omega$.

Proof. (i) and (ii) were proved above. (iii) is Theorem 4.3 (ii).

To prove (iv), which is equivalent to asserting $E_mSP>$ for any positive integer m, one has to check that if $f(u) = u^m$ and if

$$D_f(a,b) = \sum_{i,j} (f(a_i a_j + b_i b_j) - f(a_i b_j + b_i a_j)),$$

then $D_f(a,b) \geqslant 0$. As in the proof of the monotonicity conjecture and *PAS*>, let C_+ be the set of all the functions $f : (0,\infty) \to \Re$ such that $D_f \geqslant 0$. Then C_+ is a cone, closed with respect to simple convergence. We claim that if $f(u) = u^m$, m a non-negative integer, then f belongs to C_+. Indeed,

$$\begin{aligned}
&(a_i a_j + b_i b_j)^m - (a_i b_j + b_i a_j)^m \\
&\quad = \sum_{k=0}^{[\frac{m}{2}]} \binom{m}{k} \big(a_i^k a_j^k b_i^{m-k} b_j^{m-k} + a_i^{m-k} a_j^{m-k} b_i^k b_j^k - a_i^k b_j^k b_i^{m-k} a_j^{m-k} \\
&\qquad - a_i^{m-k} b_j^{m-k} b_i^k a_j^k\big) \\
&\quad = \sum_{k=0}^{[\frac{m}{2}]} \binom{m}{k} \big(a_i^k b_i^{m-k} a_j^k b_j^{m-k} + a_i^{m-k} b_i^k a_j^{m-k} b_j^k - a_i^k b_i^{m-k} b_j^k a_j^{m-k} \\
&\qquad - a_i^{m-k} b_i^k b_j^{m-k} a_j^k\big).
\end{aligned}$$

Let $S_{a,b}(k,l) = \sum_{i=1}^n a_i^k b_i^l$. Summing that over i, j we get

$$\begin{aligned}
D_f(a,b) &= \sum_{k=0}^{[\frac{m}{2}]} \binom{m}{k} \big(S_{a,b}(k, m-k)^2 + S_{a,b}(m-k, k)^2 \\
&\qquad - 2S_{a,b}(k, m-k) S_{a,b}(m-k, k)\big) \\
&= \sum_{k=0}^{[\frac{m}{2}]} \binom{m}{k} \big(S_{a,b}(k, m-k) - S_{a,b}(m-k, k)\big)^2 \\
&= \frac{1}{2} \sum_{k=0}^{m} \binom{m}{k} \big(S_{a,b}(k, m-k) - S_{a,b}(m-k, k)\big)^2
\end{aligned}$$

which proves that $D_f \geqslant 0$. As a byproduct, we see that the equality can happen only if $S_{a,b}(k, m-k) = S_{a,b}(m-k, m) = S_{b,a}(k, m-k)$ for any $0 \leqslant k \leqslant m$. $\square$

Remark. For $m = 1$, $D_f(a, b) = (S_{a,b}(0, 1) - S_{b,a}(0, 1))^2$ and the equality holds if a and b have the same sum. For $m = 2$, the equality holds if a and b have the same sum of squares. If $D_f(a, b) = 0$ for all m, then b must be a permutation of a.

Counterexample 7.2

(i) The property $E_\alpha SA = E_\alpha E_1 E_\infty$ is false if $1 < \alpha < \infty$, i.e., $(\alpha, 1, \infty) \notin \Omega$ for all $\alpha > 1$.

(ii) The property $E_\alpha SI = E_\alpha E_1 E_{-\infty}$ is false if $-\infty < \alpha < 1$. For $\alpha = 0$, this says that *PSI*> is false. Thus $(\alpha, 1, -\infty) \notin \Omega$ for all $\alpha < 1$.

(iii) The set Ω does not contain the points $(p, p, 1)$ or $(1, 1, 1/p)$ if $2 < p < \infty$.

Proof. (i) Choose $n = 3$, $a = (2, 1, 0)$, $b = (1, 0, 2)$. Then the 3×3 matrices with elements $(u_{i,j}) = (\max(a_i, a_j) + \max(b_i, b_j))$, $(v_{i,j}) = (\max(a_i, b_j) + \max(a_j, b_i))$ are

$$(u_{i,j}) = \begin{bmatrix} 3 & 3 & 4 \\ 3 & 1 & 3 \\ 4 & 3 & 2 \end{bmatrix} \quad \text{and} \quad (v_{i,j}) = \begin{bmatrix} 4 & 3 & 3 \\ 3 & 2 & 2 \\ 3 & 2 & 4 \end{bmatrix}.$$

Hence

$$\sum_{i,j} u_{i,j}^\alpha = 1 + 2^\alpha + 5(3^\alpha) + 2(4^\alpha), \sum_{i,j} v_{i,j}^\alpha = 3(2^\alpha) + 4(3^\alpha) + 2(4^\alpha),$$
$$\sum_{i,j} v_{i,j}^\alpha - \sum_{i,j} u_{i,j}^\alpha = -1 + 2(2^\alpha) - 3^\alpha = h(\alpha) \quad \text{(say)}.$$

It suffices to show that $h(\alpha) < 0$ for all $1 < \alpha < \infty$. But $h(1) = 0$. Thus it suffices to prove that $h'(\alpha) \leqslant 0$ for all $\alpha \geqslant 1$. In fact, $h'(\alpha) = (2\log 2)2^\alpha - (\log 3)3^\alpha = (\log 3)2^\alpha[(\log 4)/(\log 3) - (3/2)^\alpha] < 0$ for all $\alpha \geqslant 1$. The last inequality holds because $(\log 4)/(\log 3) < 3/2$ (since $16 < 27$).

Claim (ii) was proved in Theorem 6.7 and (iii) is Theorem 4.3(iv). $\square$

Remark. However, Ω contains the point $(1, 1, 1/2)$. So $E_1E_1E_{1/2}>$ is true but $E_1E_1E_{1/\beta}>$ is false if $2 < \beta < \infty$.

Counterexample 7.3. Generalized *AS* may fail to be true. The *PAS*> inequality may be written as

$$\sum_{i,j} \log\max(a_i + a_j, b_i + b_j) \geqslant \sum_{i,j} \log\max(a_i + b_j, a_j + b_i), \tag{7.6}$$

for $a_i > 0, b_j > 0$. One may wonder whether (7.6) can be generalized to *GAS*> by replacing $f(x) = \log x$ by an arbitrary increasing concave function f. The analogous strategy worked for *PSA* and *GAS*> holds for the concave increasing functions $f(x) = x^p, 0 < p \leqslant 1$ (Proposition 6.3). But defining $u_{i,j} = \max(a_i + a_j, b_i + b_j)$, $v_{i,j} = \max(a_i + b_j, a_j + b_i)$, unfortunately

$$\sum_{i,j} f(u_{i,j}) - \sum_{i,j} f(v_{i,j}) < 0 \tag{7.7}$$

can happen with $a, b \in [0, \infty)^n$ and a suitable choice of the increasing concave function f. For example, choose $n = 3$, $a = (6, 0, 4)$ and $b = (6, 1, 1)$. Then

$$U = (u_{i,j}) = \begin{pmatrix} 12 & 7 & 10 \\ 7 & 2 & 4 \\ 10 & 4 & 8 \end{pmatrix}, \quad V = (v_{i,j}) = \begin{pmatrix} 12 & 7 & 10 \\ 7 & 1 & 5 \\ 10 & 5 & 5 \end{pmatrix}.$$

The left side of (7.7) equals $S := [f(2) + f(8) + 2f(4)] - [f(1) + 3f(5)]$. When $f(x) = \log x$, then $S = \log(2 \times 8 \times 4^2) - \log(1 \times 5^3) = \log(256/125) > 0$ (as claimed by *PAS*>). However, if $f(x) = \min(0, x - 5)$, so that $f(x) = 0$ if $x \geqslant 5$, then $S = [f(2) + 2f(4)] - [f(1)] = [-3 + 2(-1)] - [-4] = -5 + 4 = -1 < 0$.

Now we prove that $(1, \pm\infty, t) \notin \Omega$ if $0 < t < 1$.

Proposition 7.4. *The inequalities* $SAE_{1/r}>$ *and* $SIE_{1/r}<$ *are false if* $r > 1$.

Proof. Let $r > 1$. The differences between the left side and the right side of $SAE_{1/r}$ and $SIE_{1/r}$ are, respectively,

$$\begin{aligned} D_1(a, b, r) &= \sum_{i,j} \left(a_i^{1/r} + a_j^{1/r}\right)^r \vee \left(b_i^{1/r} + b_j^{1/r}\right)^r \\ &\quad - \sum_{i,j} \left(a_i^{1/r} + b_j^{1/r}\right)^r \vee \left(b_i^{1/r} + a_j^{1/r}\right)^r, \\ D_2(a, b, r) &= \sum_{i,j} \left(a_i^{1/r} + a_j^{1/r}\right)^r \wedge \left(b_i^{1/r} + b_j^{1/r}\right)^r \\ &\quad - \sum_{i,j} \left(a_i^{1/r} + b_j^{1/r}\right)^r \wedge \left(b_i^{1/r} + a_j^{1/r}\right)^r. \end{aligned} \tag{7.8}$$

The task is to prove that there exist pairs $a, b \in [0, \infty)^n$ such that $D_1(a, b, r) < 0$ (thus contradicting $SAE_{1/r}>$) and (possibly other) pairs $a, b \in [0, \infty)^n$ such that $D_2(a, b, r) > 0$ (contradicting $SIE_{1/r}<$).

Let $x_i = a_i^{1/r}$ and $y_i = b_i^{1/r}$. Then (7.8) becomes

$$F_1(x, y, r) = \sum_{1 \leqslant i,j \leqslant n} (x_i + x_j)^r \vee (y_i + y_j)^r - \sum_{1 \leqslant i,j \leqslant n} (x_i + y_j)^r \vee (y_i + x_j)^r,$$
$$F_2(x, y, r) = \sum_{1 \leqslant i,j \leqslant n} (x_i + x_j)^r \wedge (y_i + y_j)^r - \sum_{1 \leqslant i,j \leqslant n} (x_i + y_j)^r \wedge (y_i + x_j)^r. \tag{7.9}$$

Consider the pair $x = (t, t, \dots, t)$ and $y = (1, 0, 0, \dots, 0)$. Suppose that $0 \leqslant t \leqslant 1$. Denote $F_1(x, y, r)$ by $f_1(t, r, n)$ and $F_2(x, y, r)$ by $f_2(t, r, n)$. Then

$$f_1(t, r, n) = \sum_{1 \leqslant i,j \leqslant n} (2t)^r \vee (y_i + y_j)^r - \sum_{1 \leqslant i,j \leqslant n} (t + y_j)^r \vee (y_i + t)^r,$$
$$f_2(t, r, n) = \sum_{1 \leqslant i,j \leqslant n} (2t)^r \wedge (y_i + y_j)^r - \sum_{1 \leqslant i,j \leqslant n} (t + y_j)^r \wedge (y_i + t)^r. \tag{7.10}$$

As

$$y_i + y_j = \begin{cases} 2 & \text{if } i = j = 1 \\ 1 & \text{if } i = 1, j \neq 1 \text{ or } j = 1, i \neq 1 \\ 0 & \text{elsewhere} \end{cases}$$

and

$$(t + y_i) \vee (t + y_j) = t + y_i \vee y_j = \begin{cases} t + 1 & \text{if } i = 1 \text{ or } j = 1 \\ t & \text{elsewhere} \end{cases}$$

and $t \leqslant 1$, (7.10) becomes

$$\begin{aligned} f_1(t, r, n) &= 2^r + (2n-2)(1 \vee (2t)^r) + (n-1)^2(2t)^r - (2n-1)(1+t)^r \\ &\quad - (n-1)^2 t^r, \\ f_2(t, r, n) &= (2t)^r + (2n-2)(2 \wedge (2t))^r - (t+1)^r - (n^2-1)t^r. \end{aligned} \tag{7.11}$$

We want to show that for any $r > 1$ there exist t, n such that $f_1(t, r, n) < 0$ and t', n' such that $f_2(t', r, n') > 0$. Choose $t = \frac{1}{2}$ and denote $f_1(\frac{1}{2}, r, n)$ by $g_1(r, n)$ and $f_2(\frac{1}{2}, r, n)$ by $g_2(r, n)$. So

$$\begin{aligned} g_1(r, n) &= 2^r + n^2 - 1 - (2n-1)(3/2)^r - (n-1)^2/2^r, \\ g_2(r, n) &= 2n - 1 - (3/2)^r - (n^2-1)/2^r. \end{aligned} \tag{7.12}$$

Multiplying by 2^r, our claims become

$$\begin{aligned} &\forall r > 1 \ \exists n \geqslant 2 \quad \text{such that } 4^r + 2^r(n^2-1) - (2n-1)3^r - (n-1)^2 < 0, \\ &\forall r > 1 \ \exists n \geqslant 2 \quad \text{such that } 2^r(2n-1) - n^2 - 3^r + 1 > 0, \end{aligned} \tag{7.13}$$

which can be considered inequalities of second degree in n, for a given r. Write them as

$$h_1(n) := n^2 - 2n\frac{3^r-1}{2^r-1} + \frac{3^r-1}{2^r-1} + 2^r < 0,$$
$$h_2(n) := n^2 - 2n \cdot 2^r + 2^r + 3^r - 1 < 0. \tag{7.14}$$

To prove that for any $r > 1$ there exists a positive integer $n \geqslant 2$ such that $h_j(n) < 0$, it suffices to check that there exist positive integers lying between the two roots of the equations $h_j = 0$. The two roots of $h_1 = 0$ are $n_{1,2} = \frac{3^r-1}{2^r-1} \pm \sqrt{\Delta_1}$, and of $h_2 = 0$ are $n'_{1,2} = 2^r \pm \sqrt{\Delta_2}$, where

$$\Delta_1 = \left(\frac{3^r-1}{2^r-1}\right)^2 - \frac{3^r-1}{2^r-1} - 2^r, \quad \Delta_2 = 4^r - 3^r - 2^r + 1. \tag{7.15}$$

A positive integer n between the two roots exists if and only if $\sqrt{\Delta_j} > 1/2 \Leftrightarrow \Delta_j > 1/4$. We shall prove that condition for $r \geqslant 4$ (in the first case) and for $r > 2$ (in the second case). For the remaining values of r, we shall use Lemma 8.1.

Step 1. For $1 < r < 2$, we can choose $n = 2$. In this case, (7.13) becomes

$$4^r + 3 \cdot 2^r - 3 \cdot 3^r - 1 < 0, \quad 3^r - 3 \cdot 2^r + 3 < 0, \quad \forall 1 < r < 2. \tag{7.16}$$

In the first case, $g(x) = 4^x + 3 \cdot 2^x - 3 \cdot 3^x - 1^x$ hence $m = 4$, where m refers to the notation used in Lemma 8.1. It follows from Lemma 8.1 that the equation $g = 0$ has at most $4 - 1 = 3$ solutions. But $g(0) = g(1) = g(2) = 0$. Therefore g does not change the sign on the interval $(1, 2)$. It must have the same sign as $g(3/2) = 8 + 6\sqrt{2} - 9\sqrt{3} - 1 < 0$. *In the second case*, $g(x) = 3^x - 3 \cdot 2^x + 3 \cdot 1^x$ hence $m = 3$ (in the notation of Lemma 8.1 again). By Lemma 8.1, $g = 0$ has at most two solutions. But $g(0) = 1 > 0$ and $g(1) = g(2) = 0$ imply that the roots are 1 and 2, hence the sign between the roots must be negative (as g outside the interval $(1, 2)$ is positive).

Step 2. If $2 \leqslant r \leqslant 4$, choose $n = 3$ to disprove $SAE_{1/r}$. Indeed, from (7.13), $g(x) = 4^x + 8 \cdot 2^x - 5 \cdot 3^x - 4$. To prove that $g(x) < 0$ if $2 \leqslant x \leqslant 4$, remark that $g(0) = 0$, $g(1) = 1$, $g(2) = -1$, $g(3) = -11$, $g(4) = -25$, $g(5) = 61$. So one root is $x_1 = 0$, another root x_2 satisfies $x_2 \in (1, 2)$, and the third and last one (according to Lemma 8.1) satisfies $x_3 \in (4, 5)$. It follows that the sign of g on the interval $[2, 4]$ is the sign of $g(3)$, i.e., is negative.

To disprove $SIE_{1/r}<$, we prove that $r \geqslant 2$ implies $\Delta_2 > 1/4$. Indeed, if $r \geqslant 2$, then the function $r \mapsto 4^r - 3^r - 2^r + 1 = \Delta_2(r)$ is increasing for $r \geqslant 2$ (write it as $\Delta_2(r) = 3^r([4/3]^r - [2/3]^r - 1) + 1$ and remark that both factors of the product are increasing!). Consequently, $r \geqslant 2$ implies that $\Delta_2(r) \geqslant \Delta_2(2) = 16 - 9 - 4 + 1 = 4 > 1/4$. The proof that $SIE_{1/r}<$ is false is complete.

Step 3 (only for $SAE_{1/r}$). Now $r > 4$. We shall prove that $\Delta_1 > 1/4$. Notice that $\frac{3^r-1}{2^r-1} > \left(\frac{3}{2}\right)^r > \frac{81}{16} > 5$ (as $r > 4$). On the other hand, the function $x \mapsto x^2 - x$ is increasing for $x > 1/2$; therefore

$$\left(\frac{3^r-1}{2^r-1}\right)^2 - \frac{3^r-1}{2^r-1} - 2^r > \left(\frac{3}{2}\right)^{2r} - \left(\frac{3}{2}\right)^r - 2^r = \left(\frac{9}{4}\right)^r - \left(\frac{3}{2}\right)^r - 2^r$$
$$= 2^r\left(\left(\frac{9}{8}\right)^r - \left(\frac{3}{4}\right)^r - 1\right). \tag{7.17}$$

As the function $r \mapsto (9/8)^r$ is increasing and $r \mapsto (3/4)^r$ is decreasing, their difference is increasing, too. Thus the function $r \mapsto 2^r((9/8)^r - (3/4)^r - 1)$ is increasing provided that the second factor is positive. That is true because $(9/8)^r - (3/4)^r - 1 > (9/8)^4 - (3/4)^4 - 1 = 1169/4096 > 0$. As a consequence

$$r > 4 \;\Rightarrow\; \Delta_1 > \left(\frac{9}{4}\right)^4 - \left(\frac{3}{2}\right)^4 - 2^4 = \frac{1169}{256} = 4.56640625 > 1/4. \qquad \square$$

Now we prove that for a large domain of values (q, r) the inequalities $SE_{-q}E_r{<}$ are false. In the notations of Section 2,

$$u_{i,j} = \left[(a_i^r + a_j^r)^{-\frac{q}{r}} + (b_i^r + b_j^r)^{-\frac{q}{r}}\right]^{-\frac{1}{q}} = \left[\frac{(a_i^r + a_j^r)^{\frac{q}{r}} + (b_i^r + b_j^r)^{\frac{q}{r}}}{(a_i^r + a_j^r)^{\frac{q}{r}}(b_i^r + b_j^r)^{\frac{q}{r}}}\right]^{-\frac{1}{q}}$$

therefore

$$u_{i,j} = \frac{(a_i^r + a_j^r)^{\frac{1}{r}}(b_i^r + b_j^r)^{\frac{1}{r}}}{\left((a_i^r + a_j^r)^{\frac{q}{r}} + (b_i^r + b_j^r)^{\frac{q}{r}}\right)^{\frac{1}{q}}}. \tag{7.18}$$

Similarly,

$$v_{i,j} = \frac{(a_i^r + b_j^r)^{\frac{1}{r}}(b_i^r + a_j^r)^{\frac{1}{r}}}{\left((a_i^r + b_j^r)^{\frac{q}{r}} + (b_i^r + a_j^r)^{\frac{q}{r}}\right)^{\frac{1}{q}}}. \tag{7.19}$$

Because the function $(x, y) \mapsto \frac{xy}{(x^q+y^q)^{\frac{1}{q}}}$ is continuous at 0 (for $x, y \geqslant 0$), these formulas make sense even if some of the numbers a_i, b_i equal 0.

Proposition 7.5. *If $r > (\ln 2)/\ln(2 - 2^{-1/q})$, $q > 0$, then $SE_{-q}E_r{<}$ is false and $(1, -q, r) \notin \Omega$. In the limiting case $q = 0$, for all $r > 1$, $SPE_r{<}$ is false and $(1, 0, r) \notin \Omega$.*

Proof. Let $L = \sum_{i,j} u_{i,j}$, $R = \sum_{i,j} v_{i,j}$ and $D = L - R$. A counterexample to the inequality $SE_{-q}E_r{<}$ is a pair $a, b \in [0, \infty)^n$ such that $D > 0$. Choose $n = 3$ and

$$a = (0, 1, x), \quad b = (1, x, 0). \tag{7.20}$$

As

$$u_{i,i} = \frac{2^{\frac{1}{r}} a_i b_i}{(a_i^q + b_i^q)^{\frac{1}{q}}} \quad \text{and} \quad v_{i,i} = \frac{(a_i^r + b_i^r)^{\frac{1}{r}}}{2^{\frac{1}{q}}},$$

for these a, b the matrices U and V become

$$U = \begin{pmatrix} 0 & \frac{(1+x^r)^{\frac{1}{r}}}{\left(1+(1+x^r)^{\frac{q}{r}}\right)^{\frac{1}{q}}} & \frac{x}{(1+x^q)^{\frac{1}{q}}} \\ \frac{(1+x^r)^{\frac{1}{r}}}{\left(1+(1+x^r)^{\frac{q}{r}}\right)^{\frac{1}{q}}} & \frac{2^{\frac{1}{r}}x}{(1+x^q)^{\frac{1}{q}}} & \frac{x(1+x^r)^{\frac{1}{r}}}{\left(x^q+(1+x^r)^{\frac{q}{r}}\right)^{\frac{1}{q}}} \\ \frac{x}{(1+x^q)^{\frac{1}{q}}} & \frac{x(1+x^r)^{\frac{1}{r}}}{\left(x^q+(1+x^r)^{\frac{q}{r}}\right)^{\frac{1}{q}}} & 0 \end{pmatrix},$$

$$V = \begin{pmatrix} \frac{1}{2^{\frac{1}{q}}} & \frac{2^{\frac{1}{r}}x}{\left(x^q+2^{\frac{q}{r}}\right)^{\frac{1}{q}}} & 0 \\ \frac{2^{\frac{1}{r}}x}{\left(x^q+2^{\frac{q}{r}}\right)^{\frac{1}{q}}} & \frac{(1+x^r)^{\frac{1}{r}}}{2^{\frac{1}{q}}} & \frac{2^{\frac{1}{r}}x}{\left(1+2^{\frac{q}{r}}x^q\right)^{\frac{1}{q}}} \\ 0 & \frac{2^{\frac{1}{r}}x}{\left(1+2^{\frac{q}{r}}x^q\right)^{\frac{1}{q}}} & \frac{x}{2^{\frac{1}{q}}} \end{pmatrix}.$$

Denote this particular D by $D(x)$. Suppose $r > 1$. Let $x \to \infty$. Thus $(1 + x^r)^{1/r} - x \to 0$ and

$$\begin{aligned} &u_{1,2} \to 1, \quad u_{1,3} \to 1, \quad u_{2,2} \to 2^{1/r}, \quad u_{2,3} - x/2^{1/q} \to 0, \\ &v_{1,2} \to 2^{1/r}, \quad v_{2,2} - x/2^{1/q} \to 0, \quad v_{2,3} \to 1. \end{aligned}$$

For large x we have $D(x) = 4 + 2^{1/r} - 2 - 2 \cdot 2^{1/r} - 2^{-1/q} = 2 - 2^{1/r} - 2^{-1/q} + o(x)$. Therefore

$$D(\infty) = 2 - 2^{1/r} - 2^{-1/q}.$$

For those pairs $(-q, r)$ for which $r > 1$ and $2^{1/r} + 2^{-1/q} < 2$ we have a counterexample for the inequality $SE_{-q}E_r$. The points (q, r) which are above the graph of $f(x) = (\ln 2)/\ln(2 - 2^{-1/x})$, $f : (0, \infty) \to \Re$, provide such a counterexample.

We give a different proof when $q = 0$. We use the same pair a, b. Then

$$u_{i,j} = (a_i^r + a_j^r)^{\frac{1}{r}}(b_i^r + b_j^r)^{\frac{1}{r}}, \quad v_{i,j} = (a_i^r + b_j^r)^{\frac{1}{r}}(b_i^r + a_j^r)^{\frac{1}{r}},$$

$$U = \begin{pmatrix} 0 & (1+x^r)^{\frac{1}{r}} & x \\ (1+x^r)^{\frac{1}{r}} & 2^{\frac{2}{r}}x & x(1+x^r)^{\frac{1}{r}} \\ x & x(1+x^r)^{\frac{1}{r}} & 0 \end{pmatrix},$$

$$V = \begin{pmatrix} 1 & 2^{\frac{1}{r}}x & 0 \\ 2^{\frac{1}{r}}x & (1+x^r)^{\frac{2}{r}} & 2^{\frac{1}{r}}x \\ 0 & 2^{\frac{1}{r}}x & x^2 \end{pmatrix}.$$

As $x \to \infty$, $D = L - R \to 2(2 - 2^{1/r})$. For $r > 1$, this limit is positive. □

Proposition 7.6. *If $0 < r < 1$, then the inequality SPE_r is false and $(1, 0, r) \notin \Omega$.*

Proof. The inequalities $E_pE_qE_r$ and $E_{tp}E_{tq}E_{tr}$ are equivalent if $t \neq 0$ (Theorem 2.9). So $SPE_r = E_1E_0E_r$ is equivalent to $E_{1/r}E_0E_1 < = E_pPS<$ with $p = 1/r > 1$. To disprove it, choose $a = (1, 1, \ldots, 1)$ and $b = (0, 0, \ldots, 0)$. After raising both sides of the inequality to the power p we get

$$L^p = 8^p + 2(n-1)4^p, \quad R^p = 9^p + 2(n-1)3^p + (n-1)^2. \tag{7.21}$$

To disprove $E_pPS<$, it suffices to produce for any $p > 1$ a positive integer $n \geqslant 2$ such that

$$9^p + 2(n-1)3^p + (n-1)^2 - 8^p - 2(n-1)4^p > 0. \tag{7.22}$$

The proof follows the same lines as that of Proposition 7.4: one checks that $n = 2$ is good for $1 < p < 3/2$ and that for $p > 3/2$ the discriminant of (7.22) – which is an inequality of second degree in n – is greater than 1/4. Actually $\Delta = 16^p - 2(12^p) + 8^p > 2$ if $p > 3/2$. $\square$

Remark. Propositions 7.5 and 7.6 show that the inequality $SPS<$ is an isolated case: if $p > 0, r > 0, p \neq r$, then $E_pPE_r<$ is false.

7.2. Equality

Inequalities of type *DEF* are often proved by induction on n. We claim that then it would be no loss of generality to assume that

$$\begin{pmatrix} a_i \\ b_i \end{pmatrix} \neq \begin{pmatrix} b_j \\ a_j \end{pmatrix} \quad \text{for all } i, j = 1, \ldots, n, \tag{7.23}$$

or equivalently that

$$\text{for all } i, j = 1, \ldots, n, \text{ either } a_i \neq b_j \text{ or } b_i \neq a_j. \tag{7.24}$$

When $i = j$, this means that $a_i \neq b_i$, $i = 1, \ldots, n$.

As we will show, this is a rather general phenomenon valid for many properties *DEF*. In more detail, as to E and F, we will assume that $E(x, y) = E(y, x)$ and $F(x, y) = F(y, x)$ and that a_i and b_j $(i, j = 1, \ldots, n)$ will be restricted to a given interval J. Then

$$u_{i,j} = E(F(a_i, a_j), F(b_i, b_j)), \quad v_{i,j} = E(F(a_i, b_j), F(a_j, b_i)), \quad i, j = 1, \ldots, n, \tag{7.25}$$

must be well defined and the associated $n \times n$ matrices

$$U = U(a, b) = (u_{i,j}(a, b)), \quad V = V(a, b) = (v_{i,j}(a, b))$$

are symmetric. Finally, we require that property *DEF* has the following special form.

Definitions

(i) Property *DEF* is true if DEF_n is true for all $n \geqslant 1$. For $n \geqslant 1$ fixed, DEF_n is true if and only if, for each choice of $a, b \in J^n$, and $u_{i,j} = u_{i,j}(a, b)$ and $v_{i,j} = v_{i,j}(a, b)$ as in (7.25),

$$\phi U := \sum_{i=1}^{n}\sum_{j=1}^{n}\phi(u_{i,j}) \leqslant \phi V := \sum_{i=1}^{n}\sum_{j=1}^{n}\phi(v_{i,j}). \tag{7.26}$$

Here ϕ is a fixed function independent of n. The opposite case $\phi U \geqslant \phi V$ is realized by replacing ϕ by $-\phi$.

(ii) If $\phi U = \phi V$, that is, if (7.26) holds with equality, then the pair $a, b \in J^n$ is said to be an *equality pair*.

(iii) A pair $a, b \in J^n$ of n-tuples is said to be a *special pair* if there exist $s, t \in \{1, \dots, n\}$ such that

$$\text{both } a_s = b_t \text{ and } a_t = b_s. \tag{7.27}$$

All other pairs $a, b \in J^n$ that satisfy (7.26) are said to be *non-special*.

(iv) A pair $a, b \in J^n$ has *elementary structure* if and only if $\{1, 2, \dots, n\}$ completely decomposes into (disjoint) singlets $\{r\}$ satisfying $a_r = b_r$ and pairs $\{s, t\}$ satisfying $s \neq t$ and (7.27).

Remark. One example would be $a = (2, 1, 4, 2, 1, 3)$; $b = (3, 2, 4, 1, 1, 2)$ (where $n = 6$). An analogous definition of elementary structure was employed in the proof of *PSP* (see (6.7)). There we were only interested in the ratios $c_j = a_j/b_j$ so that (7.27) takes the form $c_s c_t = 1$.

Proposition 7.7

(i) *Each pair $a, b \in J^n$ of elementary structure is an equality pair.*

(ii) *Let $n \geqslant 2$ be a fixed integer and suppose that DEF_m is true for all integers $1 \leqslant m \leqslant n-1$. Then* (7.26) *is true for each special pair $a, b \in J^n$.*

(iii) *In proving DEF_n by induction with respect to n, it suffices to show that* (7.26) *is satisfied by each non-special pair $a, b \in J^n$.*

Proof. Suppose $a, b \in J^n$ is a special pair. Thus there exist $s, t \in \{1, \dots, n\}$ (to be kept fixed) such that

$$a_s = b_t = \alpha \text{ (say)}, \quad a_t = b_s = \beta \text{ (say)}; \quad \text{if } s = t, \text{ then } a_s = b_s. \tag{7.28}$$

Then

$$\begin{aligned} u_{j,s} = u_{s,j} &= E(F(a_s, a_j), F(b_s, b_j)) = E(F(\alpha, a_j), F(\beta, b_j)) \\ &= E(F(\beta, b_j), F(\alpha, a_j)) = E(F(a_t, b_j), F(b_t, a_j)) \\ &= v_{t,j} = v_{j,t}. \end{aligned}$$

Thus

$$u_{s,i} = v_{t,i}, \quad u_{t,i} = v_{s,i}, \quad i = 1, \dots, n. \tag{7.29}$$

Let $m = n-2$ if $s \neq t$ and $m = n-1$ if $s = t$. Also let U^o and V^o denote the symmetric $m \times m$ matrix obtained from U and V, respectively, by dropping rows s and t as well as the columns s and t. Then

$$\phi V - \phi U = \phi V^o - \phi U^o. \tag{7.30}$$

The matrices U^o and V^o may be written

$$U^o = U(a^o, b^o), \quad V^o = V(a^o, b^o),$$

where $a^o \in J^m$(respectively $b^o \in J^m$) is obtained from $a \in J^n$ (respectively $b \in J^n$) by dropping the coordinates a_s and a_t (the coordinates b_r and b_s respectively). It follows from (7.30) that $a, b \in J^n$ is an equality pair if and only if $a^o, b^o \in J^m$ is an equality pair.

It is now very easy to prove (i) and (ii) by induction with respect to n. The case $n = 1$ is obvious. Now let $n \geqslant 2$ and consider a pair $a, b \in J^n$ of elementary structure. Then (7.28) holds for some choice of $s, t \in \{1, \ldots, n\}$. Since $a^o, b^o \in J^m$ obviously also has elementary structure, it follows by induction that a^o, b^o is an equality pair. Applying (7.30), it follows that also a, b is an equality pair.

As to (iii), let $a, b \in J^n$ be a given special pair. Hence, s, t can be chosen so as to satisfy (7.28). By induction, we have that $\phi(V^o) - \phi(U^o) \geqslant 0$. It then follows from (7.30) that $\phi(V) - \phi(U) \geqslant 0$, that is, a, b satisfy (7.28). $\square$

Remarks. In the previous proposition, there is no assumption about the continuity of the functions E, F and ϕ. The above proof carries over to the slightly more general case where the definition of *DEF* is replaced by

$$\sum_{i \neq j} \phi(u_{i,j}) + \sum_{i=1}^{n} \psi(u_{i,i}) \leqslant \sum_{i \neq j} \phi(v_{i,j}) + \sum_{i=1}^{n} \psi(v_{i,i}). \tag{7.31}$$

Here ϕ and ψ denote given functions such that all terms in (7.25) are well defined for each choice of $a, b \in J^n$.

7.3. Review of open questions

The main open problem is to derive more unified proofs of the inequalities.

A second major challenge is to complete the description of the set $\Omega = \{(p, q, r) \in [-\infty, \infty]^3;\ E_pE_qE_r{<}$ or $E_pE_qE_r{>}$ hold(s)$\}$. For partial results, see Theorem 2.10, Theorem 4.3, Corollary 6.5 and the following Remark, Proposition 7.1, Counterexamples 7.2, Propositions 7.4–7.6. As part of this project, it remains to prove the monotonicity conjecture (from section 4) that if $0 < p \leqslant q \leqslant r$, then $E_pE_qE_r{<}$ holds. We proved this conjecture in Theorem 4.6 when $p = q$.

A third open problem from section 5 is to prove the conjecture that if $a, b \in [0, \infty)^n$ and if inequality SE_qE_r holds, then the corresponding QE_qE_r holds, too.

8. Zeros of sums of exponential functions, and related results

We state here some results that are used several times.

Lemma 8.1. *If* $c_1, \ldots, c_m \neq 0$ *and* $a_1, \ldots, a_m$ *are real and different, then*

$$g(x) = c_1 e^{a_1 x} + c_2 e^{a_2 x} + \cdots + c_m e^{a_m x}$$

has at most $m - 1$ *zeros, even counting multiplicities.*

Proof. Proof by induction with respect to m. The assertion is obvious when $m = 2$. Let $m \geqslant 3$ be fixed. Suppose g has k real zeros, counting multiplicities. We must show that $k \leqslant m - 1$. Clearly, $h(x) := g(x)e^{-a_m x}$ also has k real zeros, counting multiplicities. Thus its derivative $h'(x) = (g'(x) - a_m g(x))e^{-a_m x}$ has at least $k - 1$ real zeros, counting multiplicities, and so does

$$H(x) := h'(x)e^{a_m x} = c_1(a_1 - a_m)e^{a_1 x} + \cdots + c_{m-1}(a_{m-1} - a_m)e^{a_{m-1} x}.$$

But H has only $m - 1$ terms. Hence, by induction, H has no more than $m - 2$ real zeros, counting multiplicities. Hence, $k - 1 \leqslant m - 2$, that is, $k \leqslant m - 1$. $\square$

In the sequel, we shall be interested in several closed cones $\mathbf{C}$ all consisting of functions $f : (0, \infty) \to \Re$. Here, "closed" is always relative to the topology of pointwise convergence.

Definition. Let $\mathbf{C_0}$ denote the closed cone generated by the functions $f_a(x) = e^{-ax}$ with $a \geqslant 0$ together with the constant functions $f(x) = c$, $c \in \Re$. Thus, $\mathbf{C_0}$ can be described as the class of all functions $f : (0, \infty) \to \Re$ equal to the limit of some pointwise convergent functions $(f_k)_k$ of the form

$$f_k(x) = c_k + \sum_{i=1}^{m_k} b_{k,i} e^{-a_{k,i} x}, \quad \text{with } c_k \in \Re, \ \ b_{k,i} \geqslant 0, \ \ a_{k,i} \geqslant 0. \tag{8.1}$$

Remarks. Obviously

$$f \in \mathbf{C_0} \ \ \Rightarrow \ \ g \in \mathbf{C_0} \quad \text{when } g(x) = f(x + a) \ \text{ for some } a \geqslant 0. \tag{8.2}$$

Let $h > 0$. Denote by Δ_h the difference operator defined by $\Delta_h f(x) = f(x + h) - f(x)$. Let the nth iterate of Δh be $\Delta_h^n = \Delta_h \circ \Delta_h \circ \ldots \circ \Delta_h$ (n times). Clearly, if $f(x) = e^{-ax}$, $a \geqslant 0$, then $(\Delta_h^n f)(x) = f(x)(e^{-ah} - 1)^n$ and if f is constant, then $(\Delta_h^n f_a)(x) = 0$. In both cases,

$$(-1)^n \Delta_h^n f \geqslant 0, \quad \forall n \geqslant 1. \tag{8.3}$$

A non-negative function with the property (8.3) is called *completely monotonic*. Due to the linearity of Δ_h, the inequality (8.3) holds for every f of the form (8.1). As Δ_h is also continuous with respect to pointwise convergence, we have proved that any function f from $\mathbf{C_0}$ satisfies (8.3), i.e., any f from $\mathbf{C_0}$ is completely monotonic. As a consequence, any f from $\mathbf{C_0}$ is non-increasing and convex.

If, in addition, f is non-negative, then Bernstein's theorem (see [6–p. 161]) asserts that f is a Laplace transform of some measure μ on $(0, \infty)$, i.e., there exists a measure μ on $(0, \infty)$ such that

$$f(x) = \int_0^\infty e^{-zx}\,\mathrm{d}\mu(z) < \infty, \quad \forall x \geqslant 0. \tag{8.4}$$

Here μ is a unique measure on $[0,\infty)$ such that $\mu([0,a]) < \infty$, $\forall a \geqslant 0$. In fact, (8.4) implies that $\mu([0,a]) \leqslant f(x)e^{ax}$, $\forall a, x > 0$. It also follows from (8.4) that f is of class C^∞ and even analytic on $(0,\infty)$, so that (8.3) is equivalent to $(-1)^n f^{(n)} \geqslant 0$. (Indeed, divide (8.3) by h^n and let $h \to 0$.)

As a further consequence, a function f that is bounded below belongs to $\mathbf{C_0}$ if and only if $f - \inf f$ is completely monotonic on $(0,\infty)$.

Proposition 8.1. *Each of the following functions* $f : (0,\infty) \to \Re$ *belongs to* $\mathbf{C_0}$:

(i) *Any finite Laplace transform of the form* (8.4).
(ii) *For all* $a \geqslant 0$ *and* $p > 0$, *the function* $f(x) = (a+x)^{-p}$. *This includes the function* $f(x) = x^{-p}$, $p \geqslant 0$.
(iii) *For all* $a \geqslant 0$, *the function* $f(x) = -\log(a+x)$.
(iv) *For each* $0 \leqslant p \leqslant 1$, *the function* $f(x) = -x^p$.

Proof. (i) The function $f(x)$ can be represented as the pointwise limit of finite sums $f_k(x) = \sum_i c_{k,i} \exp(-a_{k,i}x)$ such that $c_{k,i} \geqslant 0$, $a_{k,i} \geqslant 0$ (since the measure μ can be weakly approximated with discrete measures). Thus $f_k \in \mathbf{C_0}$.

(ii) follows immediately from the representation

$$(a+x)^{-p} = \frac{1}{\Gamma(P)} \int_0^\infty z^{p-1} e^{-z(a+x)}\,\mathrm{d}z \quad \text{if } a \geqslant 0 \text{ and } p > 0. \tag{8.5}$$

(iii) As the cone $\mathbf{C_0}$ contains the constants, by (ii) it must contain the functions $x \mapsto (a+x)^{-p} + (-1)$, for all $p > 0$. Hence it contains the functions $x \mapsto [(a+x)^{-p} - 1]/p$, $p > 0$. But $[(a+x)^{-p} - 1]/p \to -\log(a+x)$ as $p \to 0, p > 0$. The cone is closed. Thus it contains limit functions such as $x \mapsto -\log(a+x)$.

(iv) We know from (ii) with $p = 1$ that, for all $a \geqslant 0$, the function $x \mapsto 1/(a+x)$ is in $\mathbf{C_0}$. Hence so are the functions $x \mapsto a/(a+x) - 1 = -x/(a+x)$ provided $a \geqslant 0$. Hence, so are the finite-valued functions f on $(0,\infty)$ that admit a representation of the form

$$f(x) = \int_0^\infty \frac{-x}{a+x}\,\mathrm{d}\mu(a), \tag{8.6}$$

where μ is any non-negative measure on $[0,\infty)$. Suppose $0 < p < 1$. Then, choosing μ to be the measure which has the density ρ with respect to Lebesgue measure, where

$$\rho(a) = a^{p-1} 1_{(0,\infty)}(a)/B(p, 1-p),$$

one finds that $x \mapsto -x^p$ belongs to $\mathbf{C_0}$. □

Remark. By (ii), the function φ_p given by $\varphi_p(x) = x^p$ belongs to $\mathbf{C_0}$ if $p \leqslant 0$. But if $p > 0$, then $\varphi_p \notin \mathbf{C_0}$ since it is strictly increasing. Next, if $\psi_p(x) = -x^p$, then by (iv), $\psi_p \in \mathbf{C_0}$ if $0 \leqslant p \leqslant 1$. But $\psi_p \notin \mathbf{C_0}$ if $p < 0$ (since it is then strictly increasing) and also $\psi_p \notin \mathbf{C_0}$ if $p > 1$ (since ψ_p is then strictly concave).

Definition of Property (A). Let $\mathbf{C}$ be any class of functions $f : (0, \infty) \to \Re$. We say that $\mathbf{C}$ has property (A) if all of the following are true:

(i) $\mathbf{C}$ is a cone closed with respect to pointwise convergence.
(ii) For each $a \geqslant 0$, $f_a(x) = e^{-ax}$ belongs to $\mathbf{C}$.
(iii) $\mathbf{C}$ contains all the constants $f(x) = c$, $c \in \Re$ (positive or negative).

Corollary 8.2. *A class $\mathbf{C}$ of functions $f : (0, \infty) \to \Re$ has property* (A) *if and only if $\mathbf{C}$ is a closed cone and $\mathbf{C_0} \subset \mathbf{C}$. Hence, property* (A) *implies that $\mathbf{C}$ contains all the special functions* (i)–(iv) *in Proposition* 8.1.

Proposition 8.3. *Let $\mathbf{C_1}$ consist of all the functions $f : (0, \infty) \to \Re$ such that*

$$\begin{aligned} D_n f &:= \sum_{1 \leqslant i, j \leqslant n} (f(x_i + x_j) + f(y_i + y_j) - f(x_i + y_j) - f(y_i + x_j))\xi_i \xi_j \\ &\geqslant 0 \end{aligned} \tag{8.7}$$

for any $n \geqslant 1$, vectors $x, y \in (0, \infty)^n$ and $\xi \in \Re^n$. Then $\mathbf{C_1}$ has property (A).

Proof. The conditions (i) and (iii) are obviously true. Condition (ii) follows from

$$D_n f_a = \left(\sum_{1 \leqslant i \leqslant n} (e^{-ax_i} - e^{-ay_i})\xi_i \right)^2 \geqslant 0. \qquad \square \tag{8.8}$$

Proposition 8.4. *Let $\mathbf{C_2}$ consist of all the functions $f : (0, \infty) \to \Re$ such that*

$$\begin{aligned} D_n f &:= \sum_{1 \leqslant i, j \leqslant n} [-f((x_i + x_j) \vee (y_i + y_j)) + f((x_i + y_j) \vee (y_i + x_j))]\xi_i \xi_j \\ &\geqslant 0 \end{aligned} \tag{8.9}$$

for any $n \geqslant 1$ $x, y \in (0, \infty)^n$ and $\xi \in \Re^n$. Then $\mathbf{C_2}$ has property (A).

Proof. Conditions (i) and (iii) are obviously true. Condition (ii) follows from *QIP*<, which we know to be true. $\square$

Proposition 8.5. *Let $\mathbf{C_3}$ consist of all the functions $f : (0, \infty) \to \Re$ such that*

$$\begin{aligned} D_n f &:= \sum_{1 \leqslant i, j \leqslant n} [f((x_i + x_j) \wedge (y_i + y_j)) - f((x_i + y_j) \wedge (y_i + x_j))]\xi_i \xi_j \\ &\geqslant 0 \end{aligned} \tag{8.10}$$

for any $n \geqslant 1$, $x, y \in (0, \infty)^n$ and $\xi \in \Re^n$. Then $\mathbf{C_3}$ has property (A).

Proof. Conditions (i) and (iii) are obviously true. Condition (ii) follows from $QAP>$, which we know to be true. $\square$

Acknowledgments

We thank loan Cuculescu, Victor de la Peña, Daniel C. Reuman and Radu Zaharopol for helpful comments. J.E.C. acknowledges with thanks the support of US National Science Foundation grants BSR9207293 and DEB9981552 and the hospitality of Mr. and Mrs. William T. Golden during this work.

References

[1] F. Baccelli, D. Hong, Analytic expansions of max–plus Lyapunov exponents, Ann. Appl. Probab, 10 (3) (2000) 779–827.
[2] A. Buja, B.F. Logan, J.A. Reeds, L.A. Shepp, Inequalities and positive definite functions arising from a problem in multidimensional scaling, Ann. Statist. 22 (1) (1994) 406–438.
[3] J.E. Cohen, Subadditivity, generalized products of random matrices and operations research, SIAM Rev. 30 (1) (1988) 69–86.
[4] R. Cuninghame-Green, Minimax Algebra, Lecture Notes in Economics and Mathematical Systems 166, Math. Rev. 82a:90043, Springer-Verlag, 1979.
[5] F.R. Gantmacher, Theory of Matrices, Chelsea, New York, 1960.
[6] D.V. Widder, The Laplace Transform, Princeton University Press, Princeton NJ, 1946.
[7] G. Zbaganu, A new inequality with applications in measure and information theories, in: Proceedings of the Romanian Academy, Series A 1 (1), 2000, pp. 15–19.